Agonistic Interfaces: Prompting Debate And Reflection On Socio-Political Issues Using Digital Technology

Thomas Edgar Ray Feltwell

PhD

March 2020

This is dedicated to the people that
were down from day one

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Throughout the duration of my PhD I have worked as a Senior Research Assistant at Northumbria University, within the Northumbria Social Computing (NorSC) Lab. My position has been on the *Challenging online feaR And OtheRing* (CuRAtOR) project, a three year long EPSRC funded project. During the project I worked collaboratively with the research team, which consisted of Prof. Shaun Lawson (Principle Investigator), Prof. John Vines (Co-investigator), Prof. Julie Barnett (Co-investigator, University of Bath), Dr. Karen Salt (Co-investigator, University of Nottingham), Dr. Vanessa Pupavac (Co-investigator, University of Nottingham), Dr. Phillip Brooker (Senior RA, University of Bath), Dr. Gavin Wood (Senior RA), Dr. Kiel Long (Senior RA), Dr. Tom Schofield (Senior RA, University of Newcastle), and Scarlett Rowland (Research Intern).

All of the studies contained within this thesis were reviewed and approved by the Ministry of Defence Research Ethics Committee (MoDREC), as a requirement of CuRAtOR's funding. In total, the work went through two MoDREC applications: 639/MoDREC/15 and 798/MoDREC/16. Each of the MoDREC approved applications were ratified, by Northumbria Research Ethics, the final ethical approval for this work was granted by the Faculty Ethics Committee on 30/08/2018.

I declare that the word count of this thesis is 73,581 words.

Acknowledgements

To Nichola, you have been a huge help to me over the years, thank you so much for everything. Thank you to my parents, John and Carol: your support for me through my academic career has been amazing, and I am deeply grateful for all you have done for me over the years.

I am indebted to my supervision team. To my primary supervisor, Shaun: I am deeply grateful for your support, mentoring, and friendship, as both supervisor and manager! You have encouraged and enabled me to explore some fascinating research directions, and I appreciate all you have done for me in my research career. To Dave, my secondary supervisor: thank you. I have always appreciated the perspective you have brought to my work, as well as the continued japery and provocative angles you contribute. I would also like to thank Ben Kirman for all his support over the years. Indeed, there were many times during this thesis when I thought *"What would Ben Kirman do?"* To Enrique Encinas, thank you for continually pushing me to look at things in a different, often more provocative way. It is rare, in the words of Dapper Dan, that you find two people who are on the same vibratory plane as yourself.

Thanks to all on the CuRAtOR project, it was exciting and boundary-pushing research. In particular thank you to John Vines for all his support and guidance. Thank you to Gavin Wood for your input, I greatly enjoyed working with you. Thank you to Scarlett Rowland: it was always great to have your perspective. A further thank you to everyone in NorSC who has helped me over the years. Thank you to Jamie Mahoney for your proof reading expertise, and to Effie Le Moignan: your (effectively) pastoral support in the last months of my thesis were indispensable! To James Egan, thank you for always being there as an ear to listen and to chin wag with whilst playing videogames. I wrote the majority of this thesis ensconced in the beautiful landscape of the Cèvennes mountains of southern France, thank you Mum & Dad. Finally, I would like to thank Northumbria University, and the CIS Department for their funding support.

Abstract

Online discussions are affected by the design of digital technology, which can segment, polarise and divide discussion. The idea that digital technology presents a *public sphere* - a space where all can debate equally - is becoming increasingly challenged. Discussion is fragmented across digital spaces, each of which is subject to the constraints of the technology's design. As such their design, and how this affects online discussion is coming under increasing scrutiny from work within the HCI community, where this work is situated.

Using critical technical practice, this thesis documents the design and deployment of a suite of agonistically-informed technology prototypes, which I term *agonistic interfaces*. Leveraging adversarial design principles and the agonistic conception of democracy, these agonistic interfaces purposefully bring users into contact with socio-political issues, through reflection, discussion and criticality. This work examines three specific domains: i) socio-political social media ii) second-screening of reality TV and iii) conversational interfaces.

My results demonstrate the agonistic interfaces I designed were able create and sustain debate across a variety of contexts, and I describe a number of design techniques that can be used to engage criticality and reflection. The results highlight how the agonistic interfaces presented can provide a space for reflection on socio-political issues, how they can become a point for the concentration of various discourses, and they also raise questions about the moral acceptability of inclusive debate.

I present two primary contributions of this thesis. First, I identify agonism as a means of addressing existing problems around online discussion, and describe how agonistic interfaces can address this. Second, I demonstrate specific ways that agonistic interfaces can be designed to reconfigure debate and engagement, which I illustrate through the design and deployment of four technology prototypes oriented around the three domains of inquiry.

Publications

The following publications were a product of the work presented in this thesis. A detailed contribution statement for each publication is also included.

- Tom Feltwell, Gavin Wood, Kiel S. Long, Phillip Brooker, Tom Schofield, Ioannis Petridis, Julie Barnett, John Vines, and Shaun Lawson. 2017. “I’ve been manipulated!": Designing Second Screen Experiences for Critical Viewing of Reality TV. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI ’17). ACM, New York, NY, USA, 2252-2263. DOI: https://doi.org/10.1145/3025453.3025833

  Work from this paper can be found in Chapters 2, 4, Section 5.2, and in the discussion in Chapter 7.

  Contribution statement: Tom Schofield and I worked on the design of both apps. I led implementation, fieldwork, analysis and paper writing. Assisted in fieldwork by Gavin Wood, Kiel Long, Phillip Brooker and Scarlett Rowland. Comments and suggestions on paper from all co-authors.

- Tom Feltwell, Gavin Wood, Scarlett Rowland, Kiel S. Long, Chris Elsden, Phillip Brooker, John Vines, Pamela Briggs, Julie Barnett, and Shaun Lawson. 2019. Designing Second-Screening Experiences for Social Co-Selection and Critical Co-Viewing of Reality TV. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI ’19). ACM, New York, NY, USA. DOI: https://doi.org/10.1145/3290605.3300300

  Work from this paper can be found in Chapter 2, Section 5.3, and in the discussion in Chapter 7.

  Contribution statement: I led design, implementation, fieldwork, analysis and writing of paper. Assisted in fieldwork by Gavin Wood, and Scarlett Rowland. Comments and suggestions on paper from all co-authors.

- Tom Feltwell, Gavin Wood, Phillip Brooker, Scarlett Rowland, Eric P.S. Baumer, Kiel S. Long, John Vines, Julie Barnett, Shaun Lawson. 2020. Challenging Selective Exposure to Socio-Political Opinions via a Pushy Smart Home Device. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). ACM, New York, NY, USA. DOI: https://doi.org/10.1145/3313831.3376774

  Work from this paper can be found in Chapters 2, 6, and in the discussion in Chapter 7.

  Contribution statement: I led design, implementation, fieldwork, and paper writing. Comments and suggestions on paper from all co-authors.

Contents

List of Figures

List of Abbreviations

ACM	Association for Computing Machinery
AI	Artificial Intelligence
API	Application Programming Interface
AoIR	Association of Internet Researchers
BPS	British Psychological Society
CBT	Cognitive Behavioural Therapy
CUI	Conversational User Interface
CuRAtOR	Challenging online feaR And OtheRing
CTP	Critical Technical Practice
Dstl	Defence Science Technology Laboratory
EPSRC	Engineering and Physical Sciences Research Council
FG	Focus Group
GDPR	General Data Protection Regulation
HCI	Human Computer Interaction
IVR	Interactive Voice Recognition
P*x*	Participant *x*
PSOT	Positively Stockton on Tees
MoD	Ministry of Defence
NorSC	Northumbria Social Computing
RQ	Research Question
SMMO	Social Media Moral Observatory
T&C	Terms & Conditions
VUI	Voice User Interface
VR	Virtual Reality
Wk*x*	Week *x*

Author's Preface

In this thesis I write in the first person. I have followed a critical technical practice methodology, and therefore have exercised criticality on my own methods, and those of my practice disciplines (computer science, social computing, design). As such, using first person maintains clarity with respect to my own thoughts, and those that are attributed to a third party.

I am the sole author of this work. At points herein I describe collaboration with others, where I explicitly introduce the collaborators, and use the term 'we'. For example *"we sorted the codes into themes"*. As such my use of the first person avoids ambiguity between work I have conducted myself, and that performed with collaborators.

For clarity, I also use the term 'we' in explanatory passages, where the use of 'we' refers to the reader and I, for example *"What can we learn from this?"*. As such, my use of the term 'we' is used to delineate shared thoughts, either with the wider research community, or together with the reader.

Chapter 1

Introduction

Social and political (socio-political) issues are all around us, as they relate to the way our society is arranged. Some are remote, far away, almost abstract topics, whereas others are noticeable in daily life. The health of the world economy and financial regulations on banks are important socio-political issues, but they are abstract and we don't often *see* how they affect us. However, we see and experience the gentrification of our neighbourhood directly, or the influx of migrants to our city. All of these issues, abstract or not, are the result of the close inter-relation of social and political factors which influence the society we live in. This thesis embraces the idea that social and political factors are heavily interwoven, which Orum defines as: *"the social circumstances of politics, that is, to how politics both is shaped by and shapes other events in societies. Instead of treating the political arena and its actors as independent from other happenings in a society,* [political sociology] *treats that arena as intimately related to all social institutions"* [190] To illustrate how one shapes the other and vice versa: Government policy (political factors) dictate how much migrant labour is allowed into the country. Policy is the result of social factors, such as opinions towards migrants, and the need for more workers in struggling institutions. The policy, once enacted, will affect society - migrants will come to the country to work, bringing their own social factors (opinions, desires, needs), which intermingle with existing social factors, invoking new cultural phenomena or even leading to more government policy (to reduce migration, or possibly increase it).

As citizens we are exposed to, and engage with, these issues throughout daily life. The idea of the *public sphere* proposed by philosopher Habermas has com-

monly been used to understand the way that we as private individuals engage in discussion with others about socio-political issues, as a means of influencing democracy [110]. All involved in this public sphere are considered as equals, and the idea is that debate works to reach a consensus about each issue. We would commonly recognise this as *sharing views in public*, leading to debates with other members of the public, which all contributes to public opinion, that may be listened to by the state or others who can make policy. In Western democracy deliberation and debate around socio-political issues is held as a key tenet to a functioning democracy, allowing citizens to share their thoughts in order to reach consensus, share understanding about different circumstances, and facilitate the voting process [80]. Therefore society functions best when all views are represented, everyone can consider everyone else's perspectives, and a consensus can be reached. Thus, the discussion of socio-political issues is widespread, taking place on TV news, printed newspapers, online news outlets, and social media, as well as with those around us in the workplace, within the home and at cafés and bars.

1.1 The Power of Communication

Spoken words, those printed in a newspaper, or rendered on screen - these are three of the forms communication can take, and in putting together a sentence and communicating it, the words themselves, and the act of communicating, have an inherent set of *power dimensions*. These power dimensions relate to the social actors, the people doing the communicating, who have some element of control or influence that can be exerted through their communication. Those on the receiving end are subject to their control and influence, knowingly or not. Thus *"one actor within a social relationship will be in a position to carry out his own will despite resistance"* [51, quoting [247], p.53]. Put simply, those in power are able to influence the way that others understand a topic. This can be done on purpose, such as to further a political policy, or reinforce a cultural norm. Power dimensions also come into being through *everyday talk*, where everyday language talking about socio-political issues can frame topics differently, or can perpetuate previously crafted messages. In this way *"power is pluralist: it is exercised from unnumerable points, rather than from a*

single political center. It is not the possession of an elite, and it is not governed by a single overarching project." [177, p.21]

The enactment of power and influence through communication occurs in multiple ways, one of which is known as *Othering*. This is the process where a distinction is drawn, through language, between the Self, and an identified Other. The process of Othering has deep historical roots in society, as those who suffered with leprosy were cast out of society as the *Other*, based on fear of contracting the disease [88]. Othering is often used as a means for those in power, such as political organisations or the government, to oppress a group of people, by framing them as different - not like *us* [206, my italics]. One common mechanism for Othering is to utilise fear, suggesting that the Other represents a threat to “Us", based on some existing cultural fear, such as the erosion of national identity. A recent example of this is the Rohingya Muslims in Myanmar, who are persecuted by the government as their religious views are portrayed as a threat to society [197]. In this vein, Nash illustrates how power is relevant to socio-political issues, and how its application can change society: ‘*‘Some groups and actions are defined as problematic, and as in need of state control. For example, illegal migrants are generally seen as a problem in wealthy liberal-democracies, rather than as an economic benefit, and as a result they are at risk of being subjected to the force of the state. In contrast, rates of conviction for sexual assault remain low, though it is clearly illegal and the numbers of incidents reported to the police have been rising in recent years."* [177, p.37]

1.2 Power and Communication through Digital Technology

With the rise of digital technology, the power dimensions that occurred on traditional communication networks and systems have migrated online. Many communication technologies take the form of *platforms* where communication and debate take place. Examples are forums, news outlets and social media sites [239]. There is debate whether the public sphere as conceptualised by Habermas exists in online technologies [49, 109]. It is clear however that publics do form online, as the design of many online technologies and platforms is centred on facilitating user production

of content [239]. Social media platforms are a prime example, as they do not produce content themselves, but rely on users to create it themselves. Thus online platforms do not facilitate a single public sphere:

> *“other public spheres (counterpublics) emerged on message boards and chat rooms where discourse could flourish without the need for news media to mediate and transmit messages."* [49, p.3]

This scattering of public spheres across different technologies and platforms means that each public is subject to the design of the platform it resides on. Therefore the technology design can influence attitudes and behaviours at a societal level, through the power dimensions that are enabled or disabled in the design itself [51]. There are significant unanswered questions regarding how this translates into effects on real populations in real contexts [92].

The shift towards the production of content by users, rather than the traditional “top down" way that messages from the state filter down to citizens [49] means that traditional models of framing and moderation of socio-political issues by media outlets is upended [51]. This results in socio-political topics becoming ubiquitous in the social data streams users engage with when using online technology, such as social media ([131, 147]. Moreover, the nature of a near real-time global network for the open sharing of thoughts, such as Twitter, has been held up as a great positive for democracy, for example precipitating “revolution" by enabling the sharing of anti-government protests in response to a government crackdown, such as the Arab Spring [131]. However, these networks can also be manipulated to perform powerful acts of political will, such as manufacturing fake news, which has also been held up as a threat to democracy [164].

Problems of Online Publics

> *“*[The] *Internet has proven to be no less immune to control than media systems that came before - surveillance, censorship, and retaliation for speech have arguably become easier as more communication has moved online."* [49, p.3]

The above quote from Caplan & boyd represents the way that traditional power has moved to the online realm, made possible by the way online technologies are designed. Many new digital technologies are developed with a libertarian ideology underpinning them, meaning they do not wish to regulate people, nor do they wish to be regulated by governments. Therefore they often take a laissez-faire attitude towards legal regulation or moderation of content that appears on their platforms [229]. This has meant that these new digital technologies, and the capabilities they possess, have been appropriated and used to exert power, to the detriment of democracy. As an example, in the recent 2019 election in the UK, analysis of 6,000 adverts on Facebook, placed by the successful Consevatice party, were found to contain misleading claims. This is facilitated by Facebook's policy, driven by their desire to reduce government or state interference, is that political adverts should not be subject to fact checking, thus enabling the Conservatives to post misleading adverts at a critical moment before the election [203].

Through their design social media platforms are also guilty of exacerbating Othering, as these behaviours have also migrated into online publics, given their centrality to the exertion of power and influence. Research has demonstrated that platforms such as Twitter can become a central point for the Othering of already stigmatised societal groups [74], and the lack of moderation on platforms like Twitter (beyond a set of basic rules) means that open discussion spaces can be overwhelmed by unthinking, reactionary responses, severely reducing the viewpoint diversity and quality of discussion [41].

Research has shown that there are a variety of phenomena and factors that are exhibited through digital technology, which complicates the way publics form and function online. Users of social networks are prone to grouping together, seeking out people like themselves, in a process called *homophily* [156]. Just as with non-digital formats (e.g. newspapers), users online will often *selectively expose* themselves only to socio-political viewpoints that align with their own, such as choosing only to visit certain online news outlets [96]. Many digital technologies (e.g. search engines), and social media platforms integrate recommender algorithms as part of personalisating the user experience. These algorithms study a user's previous behaviour and subsequently present them with content the algorithm thinks they will be interested in. These recommender algorithms can unwittingly create a *filter bubble*,

where users are profiled and only ever see a filtered view of the system, never knowing they are looking through a filtered view [191], which can also unwittingly work to reinforce racial stereotypes, and perform Othering of groups of people [182]. Recommender algorithms are a useful technology to improve user experience, but this can come at the expense of viewpoint diversity, and rob users of the joy of serendipitously stumbling across new, exciting and provocative information [191]. Recent work has demonstrated using such algorithms to generate personalised news feeds for users can have a detrimental effect on democracy, as they can facilitate *polarisation* in the form of an *echo chamber* - where two large spheres of users are unaware of one another, and only interact with content within their spheres [95].

As discussed earlier, democracy relies on all citizens being able to see and hear what everyone else has to say, so society can reach a consensus. The above factors reduce the diversity of viewpoints that citizens can engage with online, encourage citizens *not* to engage with others and facilitate polarisation of society - at the expense of democracy. These digital platforms represent an institutionalisation of power, meaning they are a socio-political issue themselves. Nash describes this mechanism: *"The institutionalization of some social meanings rather than others makes it easier for some actors to realize their existing projects and goals, while others have to alter and adapt as best they can to new situations."* [177, p.38] This thesis is grounded in acknowledging and understanding these complex, problematic phenomena, and exploring ways that technology can be re-designed, or re-conceptualised, in order to moderate or negate some of these problems, with the aim of enriching public debate.

It is clear then that this work sits at the confluence of social and political factors, the design and development of digital technology, and the way that users interact with these systems. Therefore as Human-Computer Interaction researchers, of which I count myself, it is imperative upon us, as well as those involved in the commercial design and development of such systems, to understand the underlying mechanisms within the technology that we design, and how it impacts socio-political discussion online.

1.3 Addressing the Problems of Online Publics

The meeting of these issues presents an interesting context, in which technology can be designed and configured in specific ways to either challenge or entrench views, facilitate discussion or expose users to information. Some users of social media are already cognisant of these problems, and have attempted to address them. Activists often leverage the ability to create content themselves in an online space free from state interference, to organise social movements, such as the Occupy movement for social and economic justice [65], or the Arab Spring revolutions against oppressive governments [131]. Activists also leverage their understanding of how the platforms work in order to spread their message and gain support for their cause [84].

In the research community, there is work looking to reduce the polarisation created by recommender algorithms by introducing balance to a users news consumption, and to encourage users to reflect on the balance of news they are consuming online [94]. Others have designed systems to bring citizens together in debate during elections [137], or have proposed tools that allow users to explore the way mass media talks about specific political topics [15]. Digital technologies with a physical form have also been used by some to propose questions and prompt reflection during televised political debates [107], to collate discourse together in a radio-like fashion to create a space for reflection around energy conservation[101], or to prompt reflection about local socio-political issues by drawing on knowledge of the local environment [100].

1.3.1 Agonism and Adversarial Design

Central to much of this work is re-considering the status quo (the existing structures and practices) - in user interaction, or in technology design - as a means to consider alternative arrangements. As noted previously, the ideological underpinnings of social media platform design works to precipitate some of these problems. Thus, by re-conceptualising this underpinning, for example by using a different conception of what democracy should be, it presents an opportunity to address these problems. *Agonism* is a political theory of democracy that maintains that there will always be

diverse, often irreconcilable, differences in perspective. These differences should be debated in an *adversarial* manner, where your adversary is treated as a worthy opponent to debate with, but you remain respectful of their right to be there and express their view. It is a politics of inclusion, and cognisant that there will never be consensus, encourages respectful debate between all citizens as part of a healthy democratic process [169]. There are broadly three key tenets to the principle of agonism (each of which is described in more detail in Chapter 2):

1. Conflict and disagreement is inevitable because people will always have irreconcilable *antagonisms* which cannot be put aside. They should be treated as *agonisms*, the fuel for respectful debate.

2. The shared principle of democracy, and that pluralistic debate is a constitive part of democracy.

3. That a hegemony (the currently accepted view, the status quo) should always be allowed to be challenged.

Straddling the disciplines of computer science and design is *adversarial design*, proposed by DiSalvo [70], and is founded on the theory of agonism. Adversarial design works to bring diverse viewpoints and perspectives to attention through the medium of designed objects, most notably computational objects. Adversarial design *"strives to discover and express the elements that are consistive of social conditions"* [70, p.13], and does this through designed objects that allow the users a means to act in politically meaningful ways, or the act of using an object raises questions or makes obvious social or political situations that were previously obscure. DiSalvo describes why such designed objects are effective: *"Particularly in our contemporary culture that highly valorizes technology, they command attention because they work"* [70, p.119]. The motivation behind adversarially designed objects often means it is unlikely they will receive commercial success. Nonetheless, they maintain a power beyond raising awareness and critique, as they allow users to interact with them, and experience first-hand how social and political conditions could be rearranged.

Thus using adversarial design there is an opportunity to challenge existing discourses, or reconfigure the way they are consumed. Through the design of digital tools and technologies that are aware of the ways discourse and criticality are created and fostered, and that, through their design, create discussion spaces or prompt critical thinking within the user.

1.3.2 Agonistic Interfaces

Throughout this thesis I propose the use of the term *agonistic interfaces* to define the specific places within digital technology where users can be engaged in agonistic debate and reflection around social, and political, discourses. I am using the term *agonistic interfaces* to literally describe technology interfaces that support agonistic thought and interaction, *and* interfaces which are agonistically-informed in their design. In this way, an *agonistic interface* represents both a tangible interface which can be used to engage in agonism, and an interface whose design is informed by the principles of agonism. Mouffe, who popularised agonism [169], describes how an agonistic perspective makes it possible to acknowledge and be exposed to different viewpoints that have tensions and antagonisms with your own viewpoint. Critical to agonism is the acknowledgement that those who hold different viewpoints are not enemies to be destroyed, but rather adversaries worthy of engagement in debate and discussion. Simply put, you may disagree with what the other person is saying, but you respect their right to say it. Mouffe describes this as acknowledging the 'battle lines' of acceptability. Agonistic interfaces respond to the three tenets of agonism as follows:

Conflict & Respect As noted in this section, agonistic interfaces are tangible *interfaces*, and as such facilitate users to engage in agonistic debate. They do this by creating places for users to chat with each other, or provide content to users in contexts where they can debate with other people - such as face to face with those in their house. They place a minimum threshold for debate, that it must be respectful, but allow any view to be tabled, challenged.

Democracy & Pluralistic Debate Agonistic interfaces can be developed in response to a problematic context, such as where there is little diversity of viewpoints, or where users are unable to engage in debate around contentious issues. As such

agonistic interfaces purposefully bring a plurality of viewpoints together. They facilitate a space for equal discussion, rather than invoke which viewpoint is right or wrong, therefore engaging users in discussion containing many views, allowing them to make their own mind up.

Challenging Hegemony Agonistic interfaces allow a hegemony, the status quo, to be challenged. By engaging critically-informed methods, they are designed to identify existing hegemony and open it up for challenge. Furthermore they do not allow a hegemonic view to be encoded within the interface, and allow different, plural, perspectives to be put forward and explored. For example when codifying information they allow users to explore all different viewpoints, whether they agree with them or not, and use interaction design techniques to encourage users to reflect on all of the viewpoints they encounter. Furthermore they may remove context to strip users of their pre-existing prejudices and encourage them to make judgements on what they see before them.

In this work, I have used an adversarial design approach to produce digital prototypes that present agonistic interfaces, thus they are agonistically-informed in their design. However, an agonistic interface may equally be arrived at through other design practices that invoke the political, the critical and the agonistic, such as critical design, provided they present an interface that can be used to engage in agonism. Throughout the subsequent chapter I explore this thoroughly and explore examples that illustrate agonistic interfaces that are the product of agonistically-informed design approaches, and those from other, critically informed design approaches.

As will be seen in the following section, there are acknowledged digital technologies, and specific contexts that have a deleterious effect on democracy. There are opportunities for agonistic interfaces to be designed, either using novel objects, or by reconfiguring existing interfaces, to do the work of agonism - raising questions and revealing the previously obscured or under-represented socio-political perspectives.

1.4 Domains of Inquiry

The issues with online publics are evident across a range of digital technologies. I have identified three domains of inquiry that present opportunities for the design

of agonistic interfaces to enrich and facilitate debate around socio-political topics. I describe each domain, and the rationale for studying it, as follows:

1. **Socio-Political Social Media** Social media platforms facilitate large scale discussion of many topics, and allow for near-realtime discussions, as well as asynchronous debates spread over hours or days. Discussion on these platforms is wide ranging, with different platforms attracting and appealing to different audiences. Socio-political issues are evident on these platforms through the use of aggregation features, such as hashtags, by users and other organisations. Previous work has highlighted online *"backchannels"* [75], which are often unmoderated, and can lead to certain popular opinions dominating, as well as fear of reprisals and arguments during debates stymieing engagement. Social media platforms also contain an array of power structures that can be manipulated for the dissemination or propagation of messages that relate to socio-political issues, as previously discussed. This is a socio-political issue because unmoderated discussion spaces around contentious issues, that are susceptible to manipulation, can (potentially intentionally) impact social views, which in turn can also impact political policy. This domain offers an opportunity to explore how engagement with social media can be re-configured to introduce agonism into the discussion.

2. **Second-Screening of Reality TV**

   The act of watching TV whilst using another device, known as second-screening, has been identified by previous work (e.g. [41]) as having, in the UK, some specific contexts that are particularly problematic in terms of socio-political issues. The *reality TV* genre has been shown to be problematic, as it purposefully portrays people in exaggerated ways, in order to maximise viewing figures. The surrounding discussion (online and offline) is often uncritical, and the discourse is also often being guided by the TV media. As a genre it has also been noted that the viewers are typically uncritical [217]. This is a socio-political issue because reality TV enacts power upon the viewers, by presenting certain groups in specific ways, and these programmes are often viewed uncritically. Unchallenged it has the ability to influence large numbers of people, and as

demonstrated in previous work, online platforms do not successfully present a mechanism to rebalance the power enacted by the programme makers. Previous work exploring the activity of second-screening has demonstrated that viewers can be given activities to reconfigure their engagement with the TV program [8], or receive tailored prompts to encourage offline discussion around the contents of a political debate [107]. In this way, the technology is able to shape and structure the discourse around the TV viewing process, and thus open up a means for viewers to challenge and re-interpret the messages being received through the programme. Therefore this domain presents interesting opportunities and challenges for the re-design of how people watch TV and criticality engage with socio-political issues.

3. **Conversational Interfaces**

We are now able to interact easily using conversation with digital technologies, based on advances in machine learning and natural language processing. Conversational interfaces as a technology can be split into text-based chat interfaces and voice user interfaces - the unifying characteristic of both is their use of conversational features in the interface. Text-based chat interfaces are often used as a means of interacting with an automated system, for example contacting customer support, whereas voice user interfaces are commonly used for personal assistants built into devices, such as Apple's Siri built into all recent iPhones. This diversity of applications has led conversational interfaces to be integrated into our daily lives, and thus occur in spaces and situations where we encounter socio-political topics. Devices such as the Amazon Echo are already being used to engage with news and opinion [6], and Google exhibited the ability for its voice assistant to make telephone calls on the users behalf [242], demonstrating the power dimensions of conversational interfaces, and also raising concerns regarding the ethical use of such technology. Prior work has already shown that users seek information from their virtual assistants by asking questions, with children asking factual questions and (implicitly) trusting the response [218]. A further demonstration of their power, and the trust that users put in conversational interfaces, is a chatbot developed to guide refugees through the complex legal process of claiming

asylum [59]. Previous work has demonstrated the often polarising effects that the design of social media can have on public debate, which presents a socio-political issue around existing designs because they can heavily influence the outcome of political processes, or societal views, as discussed in Chapter 2. I propose conversational interfaces as a domain of inquiry based on the opportunities that this novel, nascent, technology presents to re-configure the way people are exposed to viewpoints, and stimulate more viewpoint-diverse debate.

1.5 Research Questions & Objectives

To guide my research process, I defined three overall research questions and a set of objectives.

1.5.1 Research Questions

There are three broad research questions for this work:

1. In what ways can interfaces invoke agonism?
2. How can agonistic interfaces be designed to create and promote critical reflection and engagement with socio-political topics?
3. What forms of discussion, reflection and criticality are evoked by agonistic interfaces?

1.5.2 Objectives

To help respond to the research questions, I have defined a set of objectives:

1. Demonstrate the motivation for challenging and reconfiguring discussion and engagement with socio-political issues.
2. Understand the modes of consumption and mechanisms of discussion around socio-political topics via a thorough literature review.

3. Critically reflect on existing technical practices and how this impacts engagement with socio-political issues

4. Examine the ethical issues presented when engaging human participants with potentially upsetting socio-political content, and how this affects the design of technology prototypes, and field work methodology.

5. Use a critical technical practice approach to design, build and deploy a suite of agonistic interfaces, focused on three relevant domains:

   (a) Examine and reconfigure mechanisms for engagement with social media streams to prompt reflection and encourage plurality views.

   (b) Support and facilitate critical viewing of reality TV to encourage agonistic debate around TV medium.

   (c) Explore exposure to diverse socio-political viewpoints through a conversational interface in the home.

6. Synthesise the findings from these deployments to present the strategies and techniques for the design of agonistic interfaces, and how this facilitates and guides debate, reflection and criticality around socio-political issues.

1.6 Method of Inquiry

I approach this work with a *critical technical practice* perspective, which engages criticality and reflexivity on the processes and practices used when designing and implementing technology. To do this, I critically reflect on the existing assumptions and design conventions within each domain of inquiry, and design, implement and deploy a suite of agonistic interfaces to understand how they affect engagement with socio-political topics. In doing this, I engaged with the principles of adversarial design, which advocates using designed computational objects in order to reveal social and political situations and futures that may be obscured by existing practices and assumptions. In some cases the artefacts designed through adversarial design may be themselves adversarial towards accepted practices in their own field, and as such I see the pairing of adversarial design with critical technical practice as a suitable

method of inquiry. Critical technical practice is concerned with critical analysis of accepted practices within the design and implementation of technology, and adversarial design provides a rationale to drive this critical process. As DiSalvo highlights, an important foundation to adversarial design is *"a clever use of computation as a medium. This relies on deep knowledge and often expertise in the manipulation of computational technology"* [70, p.124]. Combined with my background in computer science and the design of digital technologies, this approach will support the inquiry into agonistic interfaces.

In each of the identified domains, I performed an initial phase of critical reflection on the existing design practices and engagement with socio-political topics. Informed by this, I designed a prototype agonistic interface to reconfigure the engagement with socio-political topics, which I then implemented using a suitable digital technology to make a working digital prototype. This prototype agonistic interface was then given to participants, either in to be used in a lab-type setting, or to be used in their homes during their daily lives. Interviews and focus groups with those participants informed how the prototypes had been used, and in what ways the agonistic interface shaped engagements with socio-political topics. Following this process in each domain, I synthesised all of the results and reflected on how the agonistic interfaces responded to the research questions.

1.7 Contributions

This thesis tells us how we, as researchers and practitioners, can specifically design digital systems to enrich and diversify the debate, reflection and criticality that users engage in. This thesis demonstrates how to identify and build interfaces that reconfigure or change the way people engage with socio-political topics (agonistic interfaces).

Overall, I present two primary contributions of this thesis. First I identify agonism as a means of addressing existing problems around engagement, exposure and debate around socio-political topics, and describe agonistic interfaces as an approach to digital technology design to address this. Second I demonstrate how agonistic interfaces can function in specific contexts to reconfigure debate and engagement

with socio-political topics, through the design and deployment of four technology prototypes. Grounded in the existing design practices in each domain, as well as my own critical technical practice, these studies elaborate how to design and deploy agonistic interfaces, as well as contributing new knowledge to the domains of socio-political social media streams, second-screening of reality TV, and conversational interfaces.

1.8 Thesis Structure

I have introduced the context of the work, and in the next chapter I present a literature review (Chapter 2). Following this grounding, I then explain my methodology, where I also provide specific details for each of the studies conducted (Chapter 3). Each of the three domains of inquiry is presented after this: socio-politicised social media (Chapter 4), second-screening of reality TV (Chapter 5), and conversational interfaces (Chapter 6). The penultimate chapter discusses these studies in relation to the research questions and objectives (Chapter 7), and I conclude with my own reflections on future work with agonistic interfaces (Chapter 8).

Chapter 2

Background and Related Work

This literature review examines socio-political discussion, both how this occurs offline, and how it manifests online. Inquiring further, I explore the concept of the public sphere, and how communication involves elements of power, influence and Othering. I discuss how this has moved online, and the problems and phenomena resulting from this, such as the echo chamber and the filter bubble. I discuss the means to address these areas, via adversarial design, critical design, and the underpinning theory of agonism. I then present my rationale for selecting the three domains of inquiry. The remainder of this chapter is divided into sections addressing the domain specific literature. First I explore work addressing the filter bubble and echo chamber effects on social media, as well as work that aims to broaden exposure to socio-political viewpoints. Second I examine work around the television as a site for socio-political topics, second-screening practices, other work to encourage criticality and engagement with TV content, and the socio-political critiques of reality TV. Finally I detail work that explores conversational interfaces, smart home technology, and how these can be used in adversarial/critical ways.

2.1 Democracy and the Public Sphere

A key concept concerning the sharing of ideas, and the formation of debate, is the *public sphere*, initially described by the philosopher Jürgen Habermas [110]. He defines the public sphere as the social institutions that facilitate citizens to engage in rational debate about social and political topics. The origins of the public sphere are

rooted in the development of capitalism, and the public sphere becomes a means by which individuals bridge their private thoughts with the state (or equivalent), and the wider public. In it's earliest form, the public sphere developed as a place for the powerful bourgeoisie to discuss their concerns and affect change within society, however in the modern mass-media society, the public sphere is conceptualised as a space open to all citizens, where public opinion is to be formed out of rational debate. Habermas also talks of *communicative power*, which is the influence that the public sphere may exert on the state. In effect, this establishes the idea that discourse can work, via a public sphere, in a bi-directional manner, with the state being influenced by public opinion, and vice versa. The idea of the bi-directional public sphere is discussed by Hauser [116] who describes the creation of 'bottom up' viewpoints and opinions that are created by those *on the ground* in a discussion. Hauser talks about how social movements have their own public sphere, where those at the top produce 'top down' messages in the same way that governments produce messages and guidance to be disseminated and adopted. It is also common that those *everyday members* of a social movement, will create their own meaning based on, or sometimes in resistance to, the top down messages. These 'bottom up' messages, grounded in the vernacular, can then influence the leaders of a social movement. As such, we see that the public sphere is a bi-directional venue for information exchange, with meaning being created by the actors throughout.

The concept of the public sphere is based on the societal conception of democracy. Elster notes that a core principle to Habermas's conception of democracy, and thus the public sphere, is that in order to be legitimate, it *"must be the outcome of deliberation about ends among free, equal and rational agents"* [80, p.5]. This links back to the history of the public sphere, which came about as a means for wealthy capitalists and landed gentry to air their opinions among those they considered as their societal equals. The idea of deliberation as part of the democratic process is widespread in western society, and is a key part of facilitating the political process, such as to deliberate about issues when voting. Elster makes the point that *"it may not be obvious that arguing is the best way of making collective decisions"* [80, p.10] but also posits that arguing is not the definitive means for facilitating deliberative democracy, as voting or bargaining have equal measure. However, arguing, as part

of deliberation, precedes both voting and bargaining, and therefore *"in this sense, arguing is logically prior to all other modes of collective decision making"* [80, p.10].

Representative Views, Othering and Stigmatisation

Due to the complex nature of deliberation, there are a number of factors that can lead to a public sphere being unequal, by not representing all parties, and by extension, all groups in society, equally. As we can see, given the public sphere's predication on equal debate amongst others to reach opinions and decisions, this can create problems.

Elster notes that societal inequalities affect the quality of any public deliberation, such as differing levels of education, which will naturally vary, and thus means participants in the discussion are not equal [80, p.13]. Gargarella describes how full representation - exposing citizens to the full range of political viewpoints - requires us to know the preferences of all other citizens so that we can include them, something that is not easily achieved. Gargarella explores how bringing all of these viewpoints together for debate is positive for democracy, because it forces *"each person to modify his or her argument in order to make it acceptable to others. So deliberation may help impartiality by forcing people to filter out mere self-interested arguments"* [93, p.261]

As touched upon earlier, communication networks invoke *power* relationships, which are *"the relational capacity that enables a social actor to influence asymmetrically the decisions of other social actor(s) in ways that favor the empowered actor's will, interests, and values."* [51, p.10]. Thus, communications from a powerful actor in the network can be used to further a political agenda, exert influence or sway opinions. Another way that power can manifest in networks is in what Michel Foucault describes as *power-knowledge*, the conditions by which those in power control and define knowledge [88]. Thus, those in power are able to subjugate and oppress groups of people, as they are able to define them as different, as Foucault illustrates with the leper:

> *"Once leprosy had gone, and the figure of the leper was no more than a distant memory, these structures still remained. The game of exclusion would be played again, often in these same places, in an oddly similar*

fashion two or three centuries later. The role of the leper was to be played by the poor and by the vagrant, by prisoners and by the 'alienated'" [88, p.6]

Foucault is discussing the concept of *the Other*, which is contrasted with the *Self*. Mountz provides a clear definition of this process:

"By placing one's self at the centre, the 'other' always constitutes the outside, the person who is different. As a noun, therefore, the other is a person or group who is different from oneself. As a verb, other means to distinguish, label, categorize, name, identify, place and exclude those who do not fit a societal norm. [...] *'Othering' is the process that makes the other."* [170, p.328]

Using Foucault's example, the vagrant is defined as different, on the fringe of society, and framing them in this way is therefore *Othering* them from society. In the foundational feminist philosophical book The Second Sex, Simone de Beauvoir describes how the historical cultural process that has occured with groups stigmatised in society, and extends this to describe the Othering of women as part of patriarchal culture [63]. The process of Othering is thus performed by those who control society, and by extension those who exert power through cultural communication networks. The process of drawing differences between groups is essential to the formation of cultural and personal identity [112, 213], but this is distinct from Othering which has attaches negative and fearful labels onto the Other. Focusing on socio-political topics, the mechanisms by which powerful messages are propagated into the public sphere are detailed by Stokes, and who focuses on how Othering is enacted:

"Consider instances in which abstract narratives concerning a category of people, narratives tailored for political ends, are believed by people whom the narrative is about, even though their own experience would lead them to believe a different narrative" [225, p.134].

Expanding on this, they take the example of state welfare in the US. In this example, the anti-state welfare argument is intended for *"white middle-class voting public. But the narratives are inevitably heard by people whose lives are their subject"* [225,

p.135]. As a result, people in communities (recipients of welfare) adopt what Stokes refers to as a 'pseudo-identity', where they perform Othering on themselves within the community, in order to deflect a politically crafted narrative. Reporting the words of an activist, Stokes demonstrate this process:

> *"There's such a barrage of shame and blame and welfare recipient bashing in this country that some* [welfare] *recipients believe some of it* [...] *So as to have some self-esteem, women on [welfare] want to say, 'I'm not like other women on welfare"'* [225, p.135].

From this example, it is clear to see that Othering is a complex social process, and the crafting of messages, through the enactment of power, is crucial to this process.

Mass media also plays a roll in these processes, through the use of *framing* when talking about particular issues. Framing is a type of agenda setting, that is used by the mass media to talk about topics in carefully designed terms, which is often done to further a political campaign, or to position people or places in a good or bad light. Castells describes how the mass media sits between the public and the state, thus reconfiguring the public sphere, meaning debate and discussion is filtered through elites and the mass media [51]. However as the word communication implies, not all ideas and discourses are produced by those in power. Those who do not have power and influence in communication networks are still engaged with discussions and perform their own processes of interpretation, so they can understand ideas themselves, and express their own ideas as well. This is illustrated when concerning a social movement: *"when the movement's rank-and-file is invited to explain it, they often give different accounts once the leader leaves the room."* [116, p.25]. Those in the movement who do not stand on the podium and spread the movement's message have a different understanding, and thus talk about it in different terms. This known as vernacular rhetoric, or *everyday talk*, which thus provides an enriched account of socio-political discussion, providing insight on small acts of resistance (such as using irony out of politeness, in order to speak *around* a topic) and how talk and ideas may be divergent from the official recognised discourse [116]. Therefore understanding the way issues are discussed by powerful entities, such as governments and news agencies, is just as important as studying

the everyday talk of those who consume and engage with these messages, and form their own opinions.

In the following section I will discuss how the public sphere, deliberative democracy and the principles of networks of communication, and their subsequent power, translate into the online domain.

2.2 The Problems of the Online Public (Sphere)

Online publics are being given increased focus, with the development of the various online spaces (see Wright et al. for a history of this [253]). Online discourse often covers socio-political factors, and these are commonly encountered on social media as views and opinions from individuals, news outlets and public figures. As I will explain, it is debated, however, whether online social networks and socio-political debates online do constitute an online public sphere in the way that Habermas described [110]. Here, the notion of the public sphere online is challenged by technological phenomena that filter the information citizens receive, or encourage them to converse with like minded people, as I will discuss shortly. Given that the access to the public sphere, and ability to debate equally with all those involved is a basic requirement of the public sphere, it is apparent that the factors outlined in the previous section frustrate this process. In this way, the information and opinions that are the most appealing, and therefore easily identifiable by a technological algorithm as being more 'likeable', rise to the top, facilitated by the technological systems that support socio-political discourse [188]. Indeed, in a more recent work Habermas challenges the idea that online communication can be termed a public sphere, given that discussions are heavily fragmented (across platforms), and that they are most often, especially in social media's case, controlled by commercial interests:

> "[Online communication] *can undermine the censorship of authoritarian regimes that try to control and repress public opinion. In the context of liberal regimes, the rise of millions of fragmented chat rooms across the world tend instead to lead to the fragmentation of large but politically focused mass audiences into a huge number of isolated issue publics."* [109, p.423]

Based on this, there are a number of terms and ideas used to describe the way publics form online. Wright et al. posit that online social networks translate another offline tradition into the digital world, the idea of the *third space*. Originally defined by Oldenburg [187] when talking about the pre-Internet era, this is the idea of a place that is not work or home, a *third space*, that involves casual, informal conversation and debate, that touches upon the (socio-)political, and with conversational norms dictated by regulars. Oldenburg uses the coffee house or English pub as an example of such a space. Wright et al. posit that third spaces exist online, although they have transformed slightly from Oldenburg's offline conception: *"A third space is, thus, a formally non-political online discussion space where political talk can emerge"* [253, p.13]. A key part of the idea of the third space is that it involves everyday talk, which as mentioned previously, allows citizens to develop their own understanding and viewpoints on socio-political topics. I explore the idea of the third space in more depth in section 2.4.2 later in this chapter.

Another idea in line with Habermas' critique of the public sphere online, the work of Semaan et al. demonstrates that users of social media often use multiple platforms, in what they term the *"sprawling public sphere"*, allowing users to build up a diverse information stream. They also note that when civic agency is reduced on one platform, they will find an alternative platform or mechanism by which they can express themselves [219].

Focusing on online publics, a great deal of work has been done to understand the way debate, and the viewpoints being shared, manifest on social media, with previous work identifying a number of features and phenomena of online networks. Two important phenomena concerning the consumption of news and engagement with opinions have come to prominence in recent years: *the filter bubble* and *the echo chamber*. The filter bubble describes the negative effects of personalisation of digital services, such as news feeds and search engines, within the context of diversity of viewpoints, sources and opinions. Pariser, in coining the term filter bubble, outlined how such systems place serendipity at risk, and reduces *"coming into contact with* [...] *mind-blowing, preconception shattering experiences and ideas"* [191, p.45] . Likewise, an echo chamber is the phenomena where individuals are only exposed to opinions that align with their own. Garimella et al. 2018 describe the two component parts that form an echo chamber: *"the opinion that is shared, and the*

'chamber' that allows the opinion to 'echo'" [95, p.913]. In their study to characterise the echo chamber phenomenon, they note an echo chamber is only created by contentious topics, as it is not observed when the topic is not contentious. Furthermore, they highlight latent phenomena within echo chambers that stifle debate: bipartisan users, who are between two sides of a debate, often act as mediators between opinions, but as a result are less central in a community, and receive lower endorsement [95].

Two important factors concerning engagement with opinions are *homophily*, and *selective exposure*. Homophily is the tendency for likeminded people to coalesce, described by the proverb *'birds of a feather flock together'* [156], and selective exposure is the concious sourcing of opinions from specific mediums, thus excluding oppositional viewpoints [96]. Colleoni et al. [57] provide an insightful analysis of political homophily on Twitter, highlighting how the social qualities of the platform facilitate an echo chamber, but the news sharing qualities facilitate a public sphere. Garrett analysed selective exposure to opinions in news articles over a 6-week study with 700 US participants, and found they were *"more likely to look at information that reinforces their opinion"* [96, p.279]. They note there is a small aversion to opinion-challenging information, but users are nevertheless willing to engage with information that challenges their opinion. Sophr [223] provides an analysis of the interplay between filter bubbles, echo chambers and selective exposure, noting selective exposure plays a major role in the formation of political polarisation on Facebook.

The term *backfire effect* has been used to describe the rejection of oppositional viewpoints, and subsequent entrenching of one's own beliefs. Nyhan and Reifler demonstrated this with news corrections that entrenched ideologically grounded beliefs [184]. However Wood and Porter challenge that the backfire effect exists, noting rather that citizens adhere to facts, even ideologically challenging ones [252]. Exploring the backfire effect on Twitter, Bail et al. [10] studied users in the US who self-reported alignment to either the Democratic or Republican parties. Participants were asked to follow a Twitter bot that presented the opposite political viewpoint. Their results demonstrate that exposure resulted in a significant entrenching of views by Republican participants, and some increased entrenching of views for Democratic participants.

Ideological Underpinnings of the Technology Platforms

As can be seen, the technology platforms themselves are not blameless, given that they are geared towards facilitating users to share almost any type of content, with a fairly loose set of rules and guidelines restricting what can be shared. This is driven by the underlying political disposition of those in Silicon Valley developing the platforms, which is generally libertarian. The libertarian view encompasses the desire for freedom of expression, and a disdain for state regulation or involvement, and thus the business models of the Silicon Valley technology companies are developed with this in mind, as discussed by Taplin [229]. In their polemical alt.chi work, Aylett & Lawson describe these foundations of Silicon Valley culture rooted in ideals of technology utopianism, and libertarianism, that frame the user as empowered, through the technology. However the commercial interests, and how they exert power over the technology is evident:

> *"In a dazzling slight of hand powerful organisations and interests convince users that only their individualism matters, thus disempowering the powerless, and generating vast quantities of personal data that empower the powerful."* [9, p.6].

Given that social media platforms place such an emphasis on sharing - as it is the basis of their business model - the platforms themselves are taking a neutral stance, implying they do not impose any viewpoint or ideology themselves. However, whilst it is the intention to be neutral, as discussed previously, the hugely complex nature of communication and the way power relationships are enacted means this is not the case. Thus the design of the technology platforms themselves are also implicated, because *they* are the medium which controls, filters and recommends the content that users engage with. Wright & Street argue that the choices made by designers, or those who commission a space for online debate, directly influence the type of debate that will occur there [254].

2.3 Re-Framing Democratic Debate Through Agonism

In designing social media platforms as unregulated spaces, where users can freely share, the assumption is that the most rational viewpoints will prevail. However this assumes that people are able to put aside their antagonisms and engage in rational debate with one another to reach a consensus. An antagonism is an irreconcilable difference between two views, an example being the differing belief in God between Christian and Muslim faiths. These views are fundamentally conflicting, and it is thus an antagonism because no consensus could be reached. As highlighted by Mouffe, this is an inherent failure of liberal and libertarian political thought, as experience has shown that people do not put aside their antagonisms, and they remain locked in destructive friend vs. enemy debate, and do not treat their adversaries as equals [169]. As a result, and as can be seen by the many examples of bullying, abuse and heated discussion that are now contemporary facets of social media (e.g. abuse of those on Twitter [55, 74]), the result of this assumption within the design process means the potential for antagonistic debate is unacknowledged, leading to the aforementioned negative consequences. This is further exacerbated by the filter bubble, echo chamber and other phenomena that alter and change the way debate is conducted online, and are the enemies of pluralistic debate between citizens [169].

Here, I posit that adopting a different perspective on democratic debate, such as the agonistic conception of democracy, can inform the design of such digital systems in a different way, that may factor in antagonism between citizens, and structure that in a way so as to encourage equality in online publics.

2.3.1 Agonism

Agonism is the process of deliberation and discussion, which celebrates disagreement and channels it positively. Agonism conceptualises democracy with disagreement as a fundamental part, that people will have irreconcilable differences in their viewpoints, and that people are not able to simply ignore these differences to engage in 'rational' debate, thus conflict is inevitable. In starting from here, agonism is then able to position the opposite view as an adversary, to be debated with. The key to the agonistic conception of democracy is that all members of the debate are treated

as a worthy *adversary* - *"whose ideas might be fought, even fiercely, but whose right to defend those ideas is not to be questioned"* [169, p.7]. There are three heavily interlinked core tenets of agonism:

1. That conflict is a healthy, and inevitable part of respectful democratic debate.
2. The shared principle of democracy as a positive for society and facilitating a plurality of viewpoints as part of this.
3. The ability for the hegemony (the *status quo*) to be challenged.

What brings agonism together as a democratic conception is the centrality of *"the 'adversary', the opponent with whom one shares a common allegiance to the democratic principles of 'liberty and equality for all', while disagreeing about their interpretation"* [169, p.7]. Mouffe notes the distinction between traditional antagonistic conflict and agonistic conflict: *"what is important is that conflict does not take the form of an 'antagonism' (struggle between enemies) but the form of an 'agonism' (struggle between adversaries)"* [169, p.7]. Thus, agonism is the acknowledgement of an opponent as a worthy adversary for debate, not as an enemy who should be destroyed or disregarded because they do not share the same view, thereupon it is fundamentally a politics of inclusion.

Why is agonism relevant in this context? Mouffe talks about the shortcomings of the liberal political perspective, which is based on the assumption of rationality, and that rational views will prevail, and an underpinning idea of the public sphere as well. Here, there is the assumption that different viewpoints will be rationally debated, with the most persuasive, or dominant, viewpoint becoming the accepted, powerful, societal view, which Mouffe describes as the *hegemony*:

> *"Society is always the product of a series of practices that attempt to create a certain order* [...] *These are the practices that we call 'hegemonic practices'. Things could always be otherwise. Every order is predicated on the exclusion of other possibilities. A particular order is always the expression of a particular configuration of power relations"* [169, p.131].

Thus, the hegemony can also be thought of as the status quo, the accepted way of society. As I have laid out in this chapter, grounding digital technology on the liberal conception of democracy has led to problems in the way publics are manifest online. Inherently the liberal political perspective (and by extension, the libertarian, freedom of speech perspective embedded within social media) is individualistic - focused on the individual - and in this way facilitates the formation of the Other through debate, as Mouffe demonstrates: *"when the others, who up to now were considered as simply different, start to be perceived as putting into question* our *identity and threatening* our *existence."* [169, p.5]. Hence, adopting an agonistic perspective to democracy offers an opportunity to reconfigure this us/them relationship into a friend/adversary relationship, with which to engage in debate towards the commonly held principle of democracy. Mouffe advocates that the point is not to reach a consensus, as this would in effect implement another hegemony, but to encourage debate, and conflict [168].

In facilitating conflict and disagreement, agonism thus allows an existing hegemony to be challenged. In a traditional democratic debate, a challenging viewpoint would be dismissed out of hand because it is a challenging the hegemony (configuring as 'us vs them'), whereas in an agonistic debate, the challenger is respected for entering the debate, and their challenging view can be rebutted as desired by the public. Mouffe speaks to the concerns raised about the configuration of media and technology as gatekeepers of the online public: *"The media are playing an important role in the maintenance and production of hegemony, but it is something that can be challenged. Every hegemony can be challenged."* [50, p.967]. Mouffe talks of the echo chamber and similar effects that occur on social media:

> *"The problem is that - and I am not the only one to point to this - many people are not using this incredible possibility of choice. In fact, it perversely allows people to just live in their little worlds, and not being exposed anymore to the conflicting ideas that characterise the agonistic public space.* [...] *I do not think that this is at all good for democracy, because for me democracy is precisely this agonistic struggle where you are being bombarded by different views."* [50, p.968]

I approach this work with an agonistic conception of democracy as laid out by Mouffe, and therefore I work to emphasise diversity of viewpoints, and the exposure to this diversity, and see this as a positive for democracy. In this vein, Bozdag & van den Hozen discuss how the conception of democracy affects the framing of problems such as the filter bubble. They acknowledge that the filter bubble is considered a problem for all conceptions of democracy, but that it is framed differently by each conception. For agonists, filter bubbles present a problem *"because they hide or remove channels through which opposing viewpoints can clash vibrantly."* [34]. They highlight that one mechanism to break such a filter would be for example, to use targetted advertising, in effect weaponising the filter bubble, but this is only accessible by the wealthy, and thus excludes parts of the public.

Focusing now on the approaches to designing digital technology that addresses some of the concerns raised previously in this chapter, I will discuss an approach to design that adopts agonism.

2.3.2 Adversarial Design

Carl DiSalvo established the term *adversarial design*, which they describe as *"a kind of cultural production that does the work of agonism through the conceptualization and making of products and services and our experiences with them"* [70, p.2]. It is an umbrella term for work in a number of disciplines (not restricted solely to design) that work to *"construct our visual and material environments"* [70, p.2]. Adversarial design is therefore focused on creating debate and conflict, as per the agonistic perspective, and doing so through the creation of objects or experiences.

It engages with 'the political', and DiSalvo makes a distinction between *design for politics* and *political design* when discussing adversarial design. Broadly, this distinction can be defined as such:

> *"Whereas design for politics strives to provide solutions to given problems within given contexts, political design strives to discover and express the elements that are constitutive of social conditions"* [70, p.13].

Thus, design for politics is concerned with the mechanisms by which democracy functions, such as civic participation in the democratic process, the organisation of

political parties, or increasing transparency via data sharing. By contrast, political design expresses political opinions or positions, asks political questions of the audience, and can challenge existing ideas and assumption. Adversarial design is accordingly a type of political design, in that it makes political points by invoking agonism, by revealing existing power structures and opening them for debate. Discussing the *Million Dollar Blocks* project [138], which visualises on a map the cost of incarcerating criminals, DiSalvo demonstrates how adversarial design works to invoke agonism: *"*[the project] *reveals previously obscured configurations in the cycle of crime and incarceration, making them available for debate, further investigation, and as leverage positions in future actions"* [70, p.13].

In order to do the work of agonism, the adversarial design approach focuses around *revealing hegemony*, and then *reconfiguring the remainder*. Revealing hegemony channels the work discussed of Mouffe, through analysing the way hegemony is enacted in the particular area or medium of study, whilst reconfiguring the remainder concerns the production of an artifact that works to present an alternate hegemony, or brings to light the existing hegemony. At it's core, *"to claim that adversarial design does the work of agonism means that designed objects can function to prompt recognition of political issues and relations, express dissensus and enable contestational claims and arguments"* [70, p.12].

Adversarial design is not solely concerned with computation as a medium, as agonism can be invoked in a variety of mediums. However, DiSalvo notes that computation provides a main focus for the work of contemporary adversarial design, as it is a complex and powerful medium: *"To understand computation as a medium requires exploring the ways those components can be used to endow artifacts and systems with distinctive qualities."* [70, p.22]. As I have demonstrated in this chapter, the way components of a digital system are configured has an impact on the form and outcome of debate and engagement with socio-political issues. Therefore adversarial design presents an approach focused on understanding, and then reconfiguring these systems to change debate and engagement.

Within HCI there has been great interest in designing digital experiences to foster and encourage critical thinking, reflection and discussion around issues of contemporary societal concern. Baumer [14] explored the conceptual dimensions that underpin notions of *reflection* in HCI research, noting the assumption in his own

and other work that reflection is intrinsically valuable as an experience. He explored the notion of *Inquiry*, and suggested design strategies to support this, such as the designation of separate spaces for inquiry to take place.

Wright et al. advocate that 'everyday talk' (as discussed in Section 2.1 on page 21) should be a feature that is encouraged and studied on social media and other online spaces, as it builds public understanding and allows talk that may cover political topics. Their paper provides a comprehensive explication of the work looking at everyday talk [253, p.10]. There are opportunities to address socio-political topics beyond the confines of purely political spaces. Graham examined how political discussions take place on non-political forums (e.g. focused around a TV show), and found that they provided *"a space where participants took personal experiences and life lessons and bridged them to society at large, fostering a more personal and lifestyle-based form of politics"* [108, p.168]. Indeed, everyday discussion does exist in a variety of places, but encouraging and facilitating debate, through the production of artifacts focused on particular topics (i.e. the hegemony) is still necessary, because as Mansbridge points out, although everyday talk occurs naturally in discussion spaces, it is not always critical in nature [152].

Having discussed the philosophical and social background of this work and the related work in the area, I will now discuss the process of identifying relevant domains of inquiry, and the related work for each of the domains.

2.4 Chosen domains of inquiry

The issues described here play out in a variety of contexts across digital technology platforms. I have identified at least three domains where these issues are manifest in interesting ways. They are: i) socio-political social media; ii) second-screening of reality TV; and iii) conversational interfaces.

Here I would like to draw on a study of reality TV and activism that I conducted prior to this PhD that provided somewhat of a case study of the way that socio-technical systems can be utilised in powerful ways, often unintended by the designers. Here I have used this case study as a lens when examining the issues and problems within this context, in order to identify the interesting domains of inquiry

for this thesis. Of course, my motivation for selecting these three domains of inquiry is not driven solely by the influence of this case study, as I have discussed in this chapter, there are a number of issues with online publics, and therefore domain one, social media, is a natural domain of inquiry based on this. Second-screening is an interesting context given the existence of socio-politicised Twitter streams, and its situation within the home presents opportunities to engage in critical viewing. Considering this physical context of TV viewing, the home environment is a natural iteration upon this for domain three, and by moving away from the TV focus the interactions available between users and the rest of the home is an interesting domain to explore. However I would like to draw on this previous study [84] to illustrate more robustly the motivation for their selection.

The subject of this study was counter-discourse activism, and it focused on two activists who used social media and related technology to produce counter-discourses against an existing hegemonic discourse [84]. The focus for these activists was a contentious popular reality TV programme, *Benefits Street*, around which both activists were working, independently, to counter the discourse being produced by the programme.

The first activist Stephen Reid, an online activist, created a website entitled *Parasite Street*, which presented the overall message that tax avoidance and tax evasion cost far more to the economy than fraud to the state welfare systems, the latter being a core argument of Benefits Street. Stephen shared the website through his personal activist network, with the result that it was widely shared on Twitter. The associated hashtag (#parasitestreet) received much activity, for the duration the programme was being aired. Stephen also utilised *Thunderclap* [1], a crowdspeaking platform, to flood Twitter with a pre-made tweet at the beginning of the programme, which he carefully orchestrated to capture the attention of many Twitter users by using a variety of hashtags. During my analysis I also saw that frame bridging, the connection of two unrelated discourses, occurred with the #parasitestreet hashtag, which was appropriated by other groups with a similar message, and it became associated with a wider discourse around the UK parliamentary expenses scandal, and remained in use long after the Benefits Street programme stopped airing.

[1] `https://en.wikipedia.org/wiki/Thunderclap_(website)`

The second activist, Mike McGrother, an activist from Stockton-on-Tees, UK, created an activist campaign, *Positively Stockton-on-Tees* (PSOT), to counter the message of Benefits Street series two, which was being filmed in Stockton-on-Tees. PSOT used a variety of social media platforms, oriented towards local residents rather than the general public, and predominantly targetted the production company, rather than the programme itself. PSOT was organised as a series of physical events in Stockton, that were propagated using their social media accounts. They carefully configured their social media accounts so that their posts used separate hashtags and wording, making sure not to reference Benefits Street directly, or inject their message into wider public discussion threads, e.g. hashtags associated with Benefits Street. My analysis showed that PSOT was successful in spreading its message with local people, receiving coverage from local news organisations, gaining involvement from local people, as well as some responses from the production company themselves. These campaigns demonstrate how the activists used their understanding of the digital technologies to harness them for their own purposes, cleverly connecting services together or carefully selecting who would see their messages, to reach people and engage them in socio-political topics. I have used this as inspiration for the three domains of inquiry:

2.4.1 Domains of Inquiry

1. **Socio-Political Social Media**: The first domain of inquiry is social media streams, in particular focusing on those that contain socio-political topics. Beyond the simple fact that interesting socio-political discussion takes place there, as can be seen in the prior example, the Parasite Street campaign was conducted *through* a Twitter hashtag, and the study demonstrates how a deep understanding of the medium (as espoused by DiSalvo when discussing adversarial design) allowed Parasite Street to be a successful campaign: by leveraging Twitter's functionality to aggregate all tweets with the same hashtag together, and using this to inject their message in front of viewers of the programme, in a discussion space that cannot be moderated. Given this ability, and the problems highlighted with online publics discussed earlier in this

chapter, studying socio-political topics on social media is a natural first direction. Thus, this domain is oriented towards examining and reconfiguring the interaction with social media in identified problematic contexts.

2. **Second-Screening of Reality TV**: The second domain of inquiry is socio-politicised TV media and the associated discourse. Parasite Street's use of the Thunderclap to get the campaign message directly in front of live viewers who were also using Twitter whilst watching the programme highlights an interesting context to study. Second-screening, the process of interacting with a secondary screen (e.g. a smartphone) whilst viewing TV (the primary screen) therefore presents an opportunity to design systems that encourage or support criticality when viewing programmes. As demonstrated by the PSOT campaign, encouraging criticality through reflection, rather than instructing users what the problems are and why they should be critical, is a useful mechanism to support the critical process. Thus, this domain is focused on harnessing a useful mechanism to challenge an identified problematic medium (reality TV).

3. **Conversational Interfaces**: The third domain of inquiry is the conversational interface as a means of engaging with socio-political topics. Conversational interfaces do not feature explicitly in the counter-discourse activism case study, however my process for selecting this domain of inquiry was partly driven by it. Both activists demonstrated how considered use of digital technology could bring their message to a desired audience, whether that was through leveraging the algorithm of Twitter (Parasite Street), or by purposefully confining online discourse to those relevant to the local community (PSOT). As can be seen in the subsequent discussion of the literature, as a technology smart home devices and the conversational interfaces that they comprise present many opportunities to engage users in debate and reflection around socio-political topics - such as by channelling online discussion, producing prompts for discussion, or provoking a reaction from them.

These three domains are the focus of the empirical work of this thesis and I discuss the relevant literature, and the opportunities each present, on the following pages.

2.4.2 Domain One: Socio-Political Social Media

Domain one is the first entry point when exploring the issues highlighted previously in this chapter - the conception of the public sphere, how this has translated online, and how the design of platforms and the way users engage with them, has led to problems with the engagement with socio-political topics. As noted, there are contentious socio-political topics that manifest on social networks, and I will describe existing work in this area that has mapped out where these discussions take place, and worked to understand that form they take. I will also talk about work that has broadened exposure to diverse socio-political viewpoints on social media.

As I have already touched upon in this chapter, the digital technologies such as social media have transformed the way people communicate, and have facilitated a range of communication, from the mundane to the political. This has meant that *"the politics and practices that unfold are foreign (and thus terrifying) to many, while offering a new way of thinking and acting for many activists.* [126, p.177]. For example, boyd discusses how initial conceptions for the Internet was that it would produce an enlightenment, a utopia of positive sharing and openness. However, boyd points out a more realistic expectation has now come to pass, that *"mainstream people are doing mainstream things, good, bad, and ugly.* [...] *the Internet mirrors and magnifies broader social and cultural values."* [33]

Inevitably, the Internet, or more specifically, social media, has become a site for political discussion, which exists in a variety of ways. Political organisations have realised the power of social media in reaching large numbers of potential voters, as demonstrated by Barrack Obama's US Presidential campaign in 2008 hat heavily utilised social media [113]. Scholars have also looked at how political messages are crafted on social media, such as Nave et al. who note that politically active users will leverage their audience and craft their messages specifically to increase the success of their posts, such as portraying their anger or humour towards an issue [178]. This of course is a mirroring of political message crafting in the offline world, which has a long history in campaigning and the creation of propaganda [35].

Beyond the purposeful crafting and propagation of political messages on social media, as I have discussed previously, some online spaces, notably some social media, are used as third spaces, where casual, spontaneous, socio-political debate

occurs [253]. For example, van Es et al. who document the ways citizens engage in their own discussions, away from formal political processes, through Facebook groups [240].

The way that sociological and political power is enacted through these digital communication networks is discussed by boyd, who cites Castells [51]:

> "[Castells] *sees technological innovation as productive and disruptive but not necessarily as the game changer that tech utopians might envision. From his perspective, the introduction and uptake of new technologies destabilizes the status quo, forcing a scramble for power.* [...] *there's also a tidal wave happening where old systems of power are working very hard to reclaim power and clamp down on the new voices that have emerged."* [126, p.177-178]

Social media has become a site for social movements to discuss and organise, as documented recently with France's Gilet Jaunes (Yellow Vests) protests. Bristow outlines that their organisation takes place on social media, especially as a means to bypass traditional institutions such as trade unions. Of note, the Gilet Jaunes movement is marked by it's decentralised nature, as it does not have an identifiable leader. Thus messages are propagated through social networks in order to reach those who identify with the movment - Bristow even identifies how Facebook videos were used to spread ideas for the movement, including the idea to adopt the eponymous yellow high visibility vest [40]. Other national and international social movements have been studied, with a view to understanding how their messages and debate have manifest on social media, such as the Egyptian and Tunisian revolutions [147], anti-austerity movements [231] and Occupy Wall Street [65]. Buettner & Buettner provide an extensive list of 'socio-political revolutions' that have been studied in recent years [45]. However not all social movements are national/global in their reach, as studied by Crivellaro et al., local communities have also taken to platforms such as Facebook to gain support for civic projects, which invokes a wide range of community viewpoints and opinions, that are debated and reconciled [60].

One way that social media has destabilized the status quo is through a powerful aggregation mechanism, known as hashtags (#). Hashtags are user-generated

tags that are used by social networking sites to aggregate content together. Hashtags originated on Twitter, but are now found on the majority of all social media [246], and given their user generated nature, have been used to discuss and document a wide range of topics, indeed many of the social movements mentioned previously in this section utilised hashtags. They are also a place for the sharing of mundane 'everyday' content, such as by parents to document the lives of their children [142] and the food they are eating [159]. They have also been used to demonstrate participation in civic duties, as demonstrated by Mahoney et al. who studied Scottish citizens uses of Instagram to talk about the Scottish Independence Referendum of 2014 [151]. Whilst the aggregation principle behind hashtags is simple, it has had powerful implications for the way discourses are centralised, and also widely accessible. Bruns et al. posit that *publics* form ad hoc around hashtags as they become popular [43].

The user generated nature of hashtags means that alternate viewpoints or *counter* discourses can be established very easily. Giglietto & Lee analysed how a counter hashtag was created after the 2015 Charlie Hebdo terrorist attack, which presented a different viewpoint to the predominant one being propagated on Twitter and in the media more generally [105]. Ince et al. demonstrate in their analysis of #BlackLivesMatter that social media users also use hashtags to connect salient socio-political issues together, with their analysis showing the #BlackLivesMatter hashtags became a central point with which other discussions and viewpoints were attached [124]. There are other examples of communities utilising the power of hashtags to express their viewpoint, such as the #sealfie created by indigenous populations to defend their traditional way of life [117, 208]

Previously Brooker et al. were introduced, who studied the differences in topics arising over different periods of Twitter conversation centered around a contentious TV show, Benefits Street - a contentious reality TV series broadcast in the UK focusing on people claiming state welfare [41]. The programme was commonly bracketed as an example of poverty porn [127]. The authors note that when the programme was broadcast live, the viewers were driven to Twitter using hashtags displayed on screen. During live broadcast, where the majority of Twitter conversation was located, the content of tweets demonstrated an orientation to the programme as a piece of entertainment. Here, tweeters frequently commented disparagingly and

judgmentally on, for instance, 'characters' in the show, focussing on their appearance, on-screen behaviours and lifestyle. However Brooker et al. noted that during the periods between programmes a different and more critical quality of conversation was evident. This conversation explored the motivations of the producers of the programme, the juxtaposition of corporate and individual tax avoidance and welfare, and the aspects of the programme which were seemingly fictionalised, despite being presented as a documentary. Therefore overall the Twitter stream had a diversity of views, but when people were engaging with the hashtag (i.e. when watching the programme), this diversity was reduced.

Other work within the HCI community has explored political events (such as elections) and research has explored ways to reconfigure engagement with news and opinions. Kriplean et al. [137] developed a platform to encourage debate around contentious US state elections. They found, amongst others, that exposure to different opinions led users to incorporate them into their own contributions to the platform. Munson et al. [171] found that news readers could be encouraged to consume a more balanced range of news by providing feedback on the political leaning of their reading behaviours over time. In this vein, Wood et al. [251] reimagined below-the-line comments on news articles, instead promoting the use of free-hand annotation by users directly on news articles. They found it facilitated user's expressivity directly on the medium, and encouraged debate between users, exposing them to diverse viewpoints on the news articles. This could be considered an agonistic interface, because it reconfigures the accepted form of online news (it is immutable, and that comments appear below) and surfaces debate directly onto the articles themselves.

Addressing the issue of narrow exposure to socio-political viewpoints, Garimella et al. [94] proposed and evaluated a recommender algorithm that exposed opposing views to users around controversial topics. Focused on social media, they describe the 'bridging' of opposing views via sharing (retweets, shares, etc.), and their algorithm takes into account the probability of a viewpoint being accepted and subsequently shared.

2.4.3 Domain Two: Second-Screening of Reality TV

Reality TV is a broad genre encompassing programmes that focus on competitions or contrived settings (e.g. Big Brother, X Factor), to those that claim, with no small degree of contestation, to document the goings-on of a (typically marginalised) social group [78]. Reality TV is distinct from what is commonly known as a *documentary* programme, based on it's different production values, as Bignell describes:

> *"situations were devised for the purpose of shooting them, and docu-soaps like Airport (1996) which impose on real events the conventions of soap opera including editing techniques of parallel montage, character-focused narrative structure and basis in a single, geographical space and community. Looking at Reality TV in this way as a programming history that increasingly diverges from documentary results in the argument that Reality TV loses the authenticity and explanation of documentary, and develops instead towards a spectacle of the everyday that emphasizes its participants' performance of identity"* [23, p.5]

Reality TV is often upheld by producers as a type of documentary which should draw a reasoned critique from its audience. Yet it is typically presented as entertainment through it's production, as described by Bignell, being focused on contrived settings and the participant's identity. Viewer-led discussions are not typically critical or reflective but tend to comprise 'surface' commentary on people appearing in the show. Indeed, Scarborough and McCoy [217] suggested that viewers who report more moral (and by implication, more 'critical') reactions to reality TV were less likely to actually watch it. Coupled with the essentially negative portrayals reality TV deals in, the tendency for reality TV to produce uncritical viewings in the majority of its audience is problematic.

Given reality TV's primary purpose as entertainment, I am motivated by the previous work of Tremlett [236] who has noted how uncritical viewings of reality TV shows often accentuate the differences between the viewer and those on screen, leading to the entrenching of negative stereotypes and stigmatisation. In this way, reality TV can serve to undermine the lives of those that it claims to document, and problematically provides justification for their ongoing Othering and stigmatisation.

Precipitated by the effects of the 2008 global recession, combined with national implementation of austerity measures, a recent focus of debate in the UK has been a particular brand of reality TV [24] that seeks to depict people of low socio-economic status supposedly going about their everyday lives. Examples of well-known series that fit this genre, which has frequently been referred to as poverty porn [127], include Skint and Benefits Street both of which were commissioned by television broadcaster Channel 4. Despite often controversial content, this genre of TV is extremely popular with the viewing public. When broadcast in mid-week evening prime time slots, and at its peak, Benefits Street was watched by almost 6.5 million UK viewers [58] per episode. Lamb [139] documents the proliferation of this type of TV programming since 2013, which can also be casually observed in viewing figures. For instance, on Channel 5, a mainstream free-to-air channel in the UK, in September 2014 poverty porn programmes comprised two of the top 30 programmes, with 3.2 million combined viewers. In contrast, September 2016 showed 11 poverty porn programmes in the top 30, with 17.7 million combined viewers [25]. Whilst these viewing figures alone are not conclusive, they complement the work of Lamb [139] and reportage of Collier [58] examining the increase in reality TV focusing on poverty increased on the broadcast schedule.

Media, political and academic reaction to UK poverty porn TV has been varied. Right-of-centre politicians and tabloid newspapers have seized upon the popularity of the genre, as well as apparent public outrage regarding the alleged feckless and immoral behaviour of people portrayed in the shows, as a mandate for ever more punitive austerity measures and welfare reform, exemplified by Conservative MP Iain Duncan Smith who cited Benefits Street as justification for welfare reform [185]. Left-leaning commentators point out the inaccurate, simplistic and ultimately problematic framing of poverty and welfare claimants [129] - a point reinforced by the majority of scholarly work. It is argued that poverty porn TV is created to suit a right-wing neoliberal agenda that exacerbates inadequate public understanding of welfare and poverty and prevents reasoned, informed and nuanced critical or plural debate. MacDonald, for instance, exposes the falsehood perpetuated by Benefits Street that there are generations of families living in ghettos of unemployment [150]. Cole [56] observes that *'public discourse has become saturated with* [...] *pejorative stereotypes of teenage mothers, feckless fathers, troubled families and fraudulent*

claimants'. Perhaps Jensen makes the clearest link however between consumption of poverty porn and public misunderstandings of welfare, arguing that such television *"crowds out critical perspectives* [...] *making the world appear self-evident and requiring no interpretation, and creating new forms of neoliberal common-sense around welfare and social security"* [127, p.2].

This is a finding echoed in related work exploring second-screening audiences. Doughty et al. performed an analysis of the Twitter stream accompanying a reality TV programme focusing on the Irish Traveller community, and found users became disinhibited - impulsive and ignoring social norms - and participated in abuse of those on screen, which was fuelled by the Otherness of those on screen. Poignantly, they note the contrast between what is being said online in the Twitter streams, and how it would be received if reproduced in an offline format: *"the unregulated and unmoderated nature of a Twitter stream permits the posting of material which would be unacceptable and, in many cases, illegal, if spoken in public or printed for publication"* [74, p.61]

Second Screening and Critical Reflection of TV

Whilst being a site for socio-political issues, the act of TV viewing also presents a number of modes of interaction, which involves social and physical factors. Second-screening refers to the common practice of interacting with a smartphone or other device (a second screen) whilst simultaneously watching a TV broadcast (the primary screen). Content on the second screen is often unrelated to the TV programme being viewed, but it can also be directly related to the TV programme, either through the use of social media (e.g. using a show's hashtag), or with dedicated applications such as a companion app [174] designed to augment the experience. Such activities enrich the viewing experience, allowing not only the augmentation of the content on the primary screen, but also the connection of multiple viewing 'spheres' through social media and connected media [53]. Collaborative viewing, or *co-viewing*, is the process of watching a TV programme or video simultaneously with other people. Traditionally, co-viewing practices might simply have featured the co-located viewing of TV in a communal space, e.g. in homes or public viewing spaces. However, in recent years there has been extensive research and development around remote

or distributed forms of co-viewing facilitated through second-screens and social networks.

Buschow et al. [47] noted that different types of TV programmes lend themselves to different types of (second screening) discussion. For instance, tweets around TV talent shows tend to focus critically on the personnel appearing in the programme; live events foster discussions more oriented to debating the production of the programme itself; political talk shows more readily enable discussions of the political details advanced therein. Brooker et al. [41] analysed the online Twitter discussion of the UK reality TV series Benefits Street, which revealed the most prominent discussions were focused on the on-screen characters and were mostly 'knee-jerk' reactions aligned to the negative framing of the people depicted. A small proportion of the discussion was more critical, linking the content to wider socio-political issues. They suggested a design opportunity existed in the *"(re)design of social media platforms with a view to enhancing the potential to support more balanced, nuanced and reflective everyday socio-political talk"* [41]. Doughty et al. [75], along with Brooker et al. [41], have argued that there is particular value in building systems that engender more critically reflective content around reality TV.

Focusing on political news topics, Ceron et al. [52] studied Twitter users who second screen political debates in Italy. They hypothesised that compared to the main "Twittersphere", second-screening Twitter users would be trendsetters compared to the rest of Twitter. Using sentiment analysis, they confirmed that second-screeners did indeed act as trendsetters, however, their results also showed that compared to the rest of the Twittersphere, those involved in second-screening around political programmes were more likely to have more divergent views, often demonstrating opposite shifts in opinion. Their work discusses the actuality of the public sphere on social media, positing that a layered public sphere is most likely to exist. This is a powerful demonstration of how the framing of programmes can be echoed onto social media.

2.4.4 Domain Three: Conversational Interfaces

Conversational interfaces facilitate voice control over an interface, and they are commonly embedded into smartphones (Apple's Siri), or as the main interface for a

smart home device. The home presents a complex web of social rules and constraints for technology, which it is often forced to operate within. Baillie & Benyon [11] examined the role of digital technologies within the home, and amongst their findings highlight the power struggles between family members over pieces of technology, such as who a shared device belongs to, and the way a device's purpose may have different perceptions amongst the family. More specifically focused on smart home devices, Porcheron et al [195] evaluated a set of Amazon Echo devices in households and found that collective processes across members of a household emerged to control the devices, such as when the device could not understand their commands. Kirman et al. [133] approach these social dynamics as an opportunity for the design of smart home technology. They describe an embodied agent, *Nag-baztag*, that uses speech and other visual cues, unprompted by the user, to encourage the household members to conserve energy. Of note, they detail how the device verbally admonishes those who are wasteful of resources, using punishments of escalating severity, such as disabling the kettle, or switching off the freezer whilst no-one is home. In this way, we see that Nag-baztag uses "pushy" techniques, where even though the user experience can be negative, it is harnessing this pushy behaviour as a powerful means to affect behaviour change. The Nag-baztag can be considered as an agonistic interface, because it reconfigures the hegemony of the home - that the user is in control of all things - and in doing so surfaces socio-political questions, such as why individuals should feel empowered to waste energy when energy waste is an acknowledge societal problem.

Along these lines, Gaver et al. [101] describe Energy Babble, an *"automated talk-radio"* smart home device that broadcast content about energy conservation and the environment into the home. This content was automatically collated from Twitter, as well as being input by participants through a telephone-like mouthpiece on the device, and by the research team themselves. Deployed with communities interested in engaging with energy and environmental based content, the device confounded the participants' expectations, as it did not offer direct advice on energy conservation, but a selection of viewpoints and discourses. The authors posit that a device like Energy Babble extends the idea of how publics are constructed, presented by DiSalvo [69], and they do this by concentrating discourses about a specific issue into

"a focused stream that inundates listeners with the many different and potentially incompatible ways that that issue is discussed" [101, p.1124]. I would consider the Energy Babble to be an example of an agonistic interface, since it reframes discussions and information about energy conservation, drawing on a range of sources, some of which are conflicting. It thus becomes a prompt for debate and reflection for those listening to it.

In the context of smart home technology, Gorkovenko et al. [107] outlined the design and deployment of living-room situated, networked printers (named *Social Printers*) that were used as a means to facilitate, create and broadcast discussion around televised political debates. Focused on political engagement in the UK, the Social Printers produced printed paper discussion points that were created by other users in the network, as well as the research team. They highlight the way a physical device within the home environment can act as a prompt for conversations and also how it created a sense of community between the users. The physicality of the printed discussion points contributed to the reconfiguration of a mundane activity - political discussion around a television broadcast - by subverting the typical format of living room debate (as advocated by DiSalvo's Adversarial Design [70]). Hence, users of the Social Printers were confronted with a need to make sense of the new format of their information, and move from a passive to actively engaged state [237]. I would also argue that the Social Printers could be considered an agonistic interface. The existing discussion conventions in the living room - the television providing the majority of the information, and others in the room discussion - is disrupted by the Social Printers, which then injects new ideas and thoughts from other viewers.

Gaver et al. also present a personal device, the Datacatcher, which is location aware and presents socio-political content to the user throughout their daily life. It uses a variety of online sources to collate short sentences that present statistics or information about the local area, for example *"People from New Cross are in the bottom 50% of health in England."* [100, p.1598] It also allowed users to engage with short quizzes about the local area, that were based on this socio-political data. Participants found it was a powerful way to learn about the local environment, and the places that they travelled through, such as the happiness rating or the cost of housing. The Datacatcher can be seen as an agonistic interface also, because it takes an existing set of information about socio-political conditions, and channels

it through the device to combine them with location - so rather than a debate or reflection starting in a dedicated space, the device prompts reflection as users move through their environment, and receive context relevant prompts to do so. The facts are presented without explanation, leading them to speculate on what they mean about the socio-political conditions around them.

2.5 Summary

In this chapter I have reviewed the literature relevant to the thesis, initially focusing on the way democratic debate takes place in society via the public sphere, how networks for communication involve power relationships that allow the influence of others in society, and how socio-political discussions take place through deliberation and everyday talk. Exploring the idea of power further, I discussed how power-inequalities make public debate unequal, how discourse works to stigmatise groups of people and perform Othering on them. Bringing the focus towards digital technologies, I detailed how the idea of the public sphere does not translate to online technology, and that there are other conceptions of how publics form online, such as by creating third spaces, or over a sprawling public sphere. I then presented existing work that explores the role technology plays in the formation and influence of online publics, such as the filter bubble, the echo chamber, and network homophily.

Following this understanding of online publics, I discussed the opportunities to tackle these issues, guided by an understanding of the ideological underpinning of many digital systems, and how they conceptualise democratic debate. To address these problems, I then presented agonism as an alternative conception of democratic debate, which focuses on inclusivity, the ability to challenge the status quo, and which holds that conflicting viewpoints are fundamental to a functioning democracy. Further demonstrating the opportunities to address the problems with online publics, I then discussed adversarial design, an approach for the creation of objects and experiences that invoke agonism. Adversarial design encourages reflection by the user, and I have documented existing work in the HCI field concerned with reflection, and how this can be facilitated by supporting inquiry, and everyday talk.

Based on this discussion of the opportunities address problems with online publics and socio-political issues, I presented a case study of activism around reality TV,

which I used as inspiration for identifying useful domains of inquiry for this thesis. The remainder of the chapter focused on the literature relevant to the chosen domains: i) socio-political social media, ii) second-screening of reality TV, and iii) conversational interfaces.

In discussing domain one, I presented the various ways that socio-political topics are discussed online, by activists leveraging social media, and how users coalesce around central points, such as hashtags, for culturally important issues. I then demonstrated how these spaces can actualise Othering and stigmatisation, based on the form of socio-political discussion. In domain two I detailed how problematic genres of TV have an unaddressed need for criticality by their audiences. This was followed by a discussion of second-screening, and the opportunities presented by the combination of socio-political issues and second-screening activities. In domain three I detailed the opportunities presented by smart home devices and conversational interfaces, in particular their location within the complex social environment of the home. I also detailed prior work that uses smart home devices to engage users in reflection on energy conservation or televised political debates. Throughout these three domain oriented chapters, I illustrated how previous works could be seen as agonistic interfaces.

Chapter 3

Methodology

I have approached this work using *critical technical practice*, which provides a perspective where I engage criticality and self-reflection of the processes and practices of designing and implementing socio-technical systems. Specifically, I have used adversarial design to inform my critical practice, and my design practice. Adversarial design has allowed me to reveal the socio-political issues evident in existing designs and interfaces, and propose new interfaces - *agonistic interfaces* - that do the work of agonism and bring out socio-political topics for debate. To better understand how these agonistic interfaces work, I created functional digital prototypes and deployed them with participants in authentic settings, as part of a *"research in the wild"* approach, in their own homes, using their own technology. Studying the agonistic interfaces in an such settings reveals the way agonistic debate can be encouraged, guided or constrained, and how that feeds into, or is influenced by, the participant's context.

In this chapter I provide more detail on the overall methodological approach taken in answering the research questions. I document the specific methodological choices for the data collection and evaluation for the studies that comprise this work, and I conclude by discussing the ethical issues anticipated and encountered during this work.

3.1 Critical Technical Practice

Critical Technical Practice (CTP) is an approach to computer science research and technology design that involves critical theory, to reflect on existing practices, how they came to be, and how revealing this can inform future design decision making and practice. CTP was proposed by Philip Agre [4] as a means to reflect on the assumptions and constraints of computer science. Agre describes that he situated himself as an artificial intelligence (AI) researcher, and in his works about CTP he discusses the history of AI research, and how historical configurations of funding bodies and institutional practices had a powerful impact on the type of research conducted around AI, and how ideas are conceptualised. In illuminating these historical connections and assumptions existing at a disciplinary level, he espoused that such self-reflexivity allows the practitioner to re-conceptualise and reconfigure aspects of their discipline to achieve different and novel outcomes.

Agre's original discussion focused on artificial intelligence, but the idea of CTP has been adopted by other strands of computer science, notably HCI. Sengers et al. describe it as asking questions about *"what values, attitudes, and ways of looking at the world are we unconsciously building into our technology, and what are their effects?"* [220, p.49]. They lay out a set of examples from HCI where assumptions have guided practice in the discipline, and inadvertently blocked off design opportunities or approaches. They note that *"in each of these cases, researchers identified values, practices, and experiences that were unconsciously, but systematically, left out of HCI."* [220, p.49]

The *practice* aspect of CTP is vital to the approach, as assumptions and unacknowledged biases are often not discovered until the practice work is underway, as illustrated by Boehner et al.: *"the values to be espoused in system design are not necessarily easy to articulate before design has begun, but instead may emerge through a process of engagement with users, materials, and fellow researchers"* [29]. Put simply, CTP encourages making and thinking through doing. This is in contrast to methods such as user-centred design (UCD), which uses methods to inform the design of an object, and evaluates the object with users. What CTP offers here is reflexivity throughout the process, on whether a process is suitable or could be changed to yield different, unacknowledged outcomes.Therefore it challenges the

designer & maker to think about the assumptions *they* are making and how these are embedded in a design.

Bowen [2007] talks about the differences between critical design (as presented by Dunne & Raby), and CTP:

> *"In this respect they* [Dunne & Raby] *are perhaps closer to art objects than designed products - they are intended to make us think.* [...] *Mainstream design is primarily about offering answers, Dunne & Raby suggest that design should now be about asking questions.* [...] *Agre's critical technical practice produces AI systems that better represent real world phenomena. Sengers' reflective design produces HCI systems that offer a more holistic satisfaction of needs and afford some critical reflection of their use.* [...] *They are primarily about producing answers"* [32].

I am seeking to answer my research questions, given the exploratory context, CTP therefore provides a line of inquiry to explore these questions, and to do so in a way where I can examine existing design assumptions and conventions - how things are designed currently to achieve a certain goal - and re-conceptualise and reconfigure the assumptions within a technology design. Engaging in reflexivity on my own, and the wider discipline's practice, allows me to understand how designing for a concept, such as agonism, change or effect the technology and/or interaction experienced by users.

Also, CTP requires a practitioner to be interdisciplinary, to straddle their practice discipline as well as critical theory: *"A critical technical practice will, at least for the foreseeable future, require a split identity – one foot planted in the craft work of design and the other foot planted in the reflexive work of critique."* [3]. HCI is an interdisciplinary field, drawing from computer science, social science, psychology and design, and it is encouraging to see that CTP is already in use in HCI (as demonstrated by Sengers et al. [220], Boehner et al. [29] and Bowen [32] discussed on the previous pages).

Another factor that makes CTP suitable is the socio-political context of the work. As discussed in Chapter 2, there are a variety of conceptions of democracy, and in particular, if and how people should engage as part of the deliberative democracy [80]. Contentious socio-political topics are often avoided, given the strong emotional

reaction they can elicit from people, and as such some people chose to avoid such discussions. However as discussed in Section 2.2, the combination of this avoidance with technology that is designed to present enjoyable content that encourages continued engagement, creates a problem with engagement with criticality and diverse opinions. In part this problem could be exacerbated by following a UCD or participatory design methodology - where users are involved in the design process, and would therefore be unlikely to suggest designs that make them uncomfortable or confront their political views, but would instead be more likely to suggest designs that they would like to use, that are enjoyable and do not challenge them. Therefore taking a CTP approach allows me to examine the philosophical underpinnings of existing design processes, and change it to promote a new, critically informed design. In this way, *"systems may be designed, not to do what users want, but to introduce users to new, critically-informed ways of looking at the world around them"* [29].

3.1.1 *Doing* Critical Technical Practice

In taking a CTP approach, I engaged with the principles of adversarial design, which advocates using designed computational objects in order to reveal social and political sitations and futures that may be obscured by existing practices and assumptions. In some cases the artefacts designed through adversarial design may be themselves adversarial towards accepted practices in their own field, and as such I see the pairing of adversarial design with critical technical practice as a suitable method of inquiry. Critical technical practice is concerned with critical analysis of accepted practices within the design and implementation of technology, and adversarial design provides a rationale to drive this critical process. As DiSalvo highlights, an important foundation to adversarial design is *"a clever use of computation as a medium. This relies on deep knowledge and often expertise in the manipulation of computational technology"* [70, p.124]. Combined with my background in computer science and the design of socio-technical systems, this approach supports inquiry into agonistic interfaces.

As a starting point, in each of the identified domains, I performed an initial phase of critical reflection on the existing design practices and engagement with socio-political topics. Informed by this, I designed a prototype agonistic interface to recon-

figure the engagement with socio-political topics, informed by the critical reflection on existing practices. I then implemented the design using suitable digital technology into a working digital prototype. This prototype agonistic interface was then given to participants, either to be used in a lab-type setting, or to be used in their homes during their daily lives. Interviews and focus groups with those participants informed how they had used the prototypes, and in what ways they felt differently about engagements with socio-political topics in the relevant domain. Following this process in each domain, another critical reflection was performed, to understand how the use of prototype agonistic interfaces fit with wider existing practices.

It is important to acknowledge the critiques of designing technology as a means to *solve problems*. This critique is most notably voiced by Evgeny Morozov, who uses the term *technology solutionism* [167], which describes those who unthinkingly apply technology as a solution to a problem, without giving thought to why the problem exists, and whether it could be solved without technology. I see that combining CTP with adversarial design as a means to exercise reflexivity, on how and why technology is being applied in certain ways, and what, if anything, can be achieved by reconfiguring or implementing novel technology. I am considering the socio-political aspects of the work, such as motivations by companies to implement technology in specific ways, as well as those from the surrounding disciplines, that identify where the problems are manifesting themselves in the technology. Therefore I am negotiating solutionism by trying to solve real, acknowledged problems. To address these problems, I am not necessarily suggesting a new technology where maybe one was absent, but rather I present ways to reconfigure an interaction so that people engage differently with the topic at hand. My work is therefore about positioning people into a better environment that is conducive to agonism using technology, rather than presenting ways for technology to solve an acknowledged problem.

3.1.2 Digital Prototyping

A prototype is an early instance or object that can be used as a base for future designs. It is a common practice within design-oriented fields to produce a prototype in order to evaluate how it is used or to test the feasibility of an idea. Within interaction design, Moggridge broadly defines a prototype as *"A representation of a*

design, made before the final solution exists." [165, p.685]. As noted by Moggridge, prototypes can be used as a way to test out ideas sooner, rather than waiting for the ideas to be tested in a final product. And to support this, a medium that can be easily manipulated allows for rapid prototyping and testing of ideas [165, p.684-685].

Buchenau & Suri established a model of *experience prototyping*, which looks not just at the direct usage of the prototype itself, but how it fits into the user's life, their lived experience, and the social and cultural relationships within that. They define experience prototyping as:

> *"the experiential aspect of whatever representations are needed to successfully (re)live or convey an experience with a product, space or system. So, for an operational definition we can say an Experience Prototype is any kind of representation, in any medium, that is designed to understand, explore or communicate what it might be like to engage with the product, space or system we are designing"* [44, p.424-425].

Particularly, they note experience prototypes can be used in different ways, depending on what the designer would like to find out. They may be used as i) a way to understand existing user experiences; ii) to explore or evaluate design ideas and concepts; and iii) to communicate ideas. Each of these uses is unpacked as follows:

- Experience prototyping can be used as a means to *identify design opportunities* and existing design practices. It is especially useful where existing designs are unfeasible (e.g. they are too expensive to produce, do not exist yet), and is a means for designers to understand the context, and experience of the real user experience. They give an example of a design team tasked with designing a pilots interface for a remote operated submarine. In order to explore the existing user experience, they built an experience prototype in the form of a game, grounded in the operating conditions experienced by remote submarine pilots from literature. It used a camera connected to a remote TV, and two players tasked with finding an object in a room containing multiple chairs. Through playing the game, they established multiple tensions and issues that must be monitored by the pilot, and an insight into how and why certain feedback was

required by the pilot. Using experience prototypes in this way allows the designer to understand some of the more subtle user experience qualities that may have been omitted or downplayed in literature.

- Experience prototypes can be used as a means to *explore different design ideas*, by having users engage with prototypes, and studying how the user interaction manifests. In this application, users may be participants external from the design process, or the designers and/or research team themselves. The focus here is the use of prototypes to quickly explore different designs, and how these influence the user experience. As such, materials may be used that do not resemble the final design, but none-the-less can be used to explore the what it is like to use such a design. A focus here is on reconfigurability and iteration, as the materials used represent the design, but during their use they may be iterated upon by the designer.

- A further, more *communicative*, use for experience prototypes is presented by Buchenau & Suri: the use of prototypes to let a *"user understand the subjective value of a design idea by directly experiencing it"* [44, p.429]. Here the prototypes are persuasive, produced as a means to demonstrate that the new design is better or challenges an existing one. Focusing on physical prototypes, Buchenau document examples where experience prototypes can be used to communicate the "look and feel" of the design, but that it may not resemble the finished product at all. However, in using prototypes in this way, the focus is not on the object directly, but on the experience the design invokes. As such they can be used to provide a rich, hands on experience of what might be enabled by a particular design [44, p.429].

Throughout my work, I have approached the design of prototypes as a means to *explore different ideas* and to be *communicative*. For example, in Chapter 6, I prototyped new functionality in smart home assistants, with a goal of reconfiguring the mode of consumption for online socio-political opinions. And in Chapter 4 and 5 I used the prototype design as a means to communicate an alternate, more considered way to interpret socio-political opinions.

Given the importance of the naturalistic setting on the usage of a prototype, it is therefore important to consider the methodological approach to getting the designed prototypes into the hands of participants, in a way that is meaningful to the aims of this work.

3.1.3 Deploying Digital Prototypes

By giving participants a *hands on* experience with a digital prototype they are also able to use them in a variety of settings, including those outside a laboratory setting. In doing so, their experience of the prototype is grounded in a more natural setting. Allowing a participant to use a digital prototypes in their own daily context is a further important step to capturing the full richness of how the prototype interacts with their existing social relationships, time pressures and habits. Buchenau & Suri place emphasis on the fidelity of the prototype, in as much that it must be usable: *"A true Experience Prototype for users - providing a really relevant experience - seems to require a level of resolution and functionality such that it can be* "let loose" *into an everyday context and more fully integrated into people's lives"* [44, p.429]. This resonates with DiSalvo's discussion of adversarial design, where they note that computation as a medium for adversarial design is powerful, as our society *"highly valorizes technology, they command attention because they work"* [70, p.119]. Thus, deploying digital prototypes with participants in naturalistic settings can yield valuable insights.

Deploying technology prototypes outside the laboratory is often know as *research in-the-wild*, a term popularised in 2011-12 by Yvonne Rogers [209], the UK's Engineering and Physical Sciences Research Council [81], and Chamberlain et al. [54]. Research in the wild represents a paradigm for HCI research focused on situated deployment and evaluation of technology interventions with people, in their homes and throughout their daily lives. This involves studying users in the places where their interactions normally take place, such as in their home, at their place of work or on the move. It can also involve giving prototypes to participants, for them to use in a natural setting. As a method, it accommodates all of the factors that exist within normal settings - interruptions, changes in attention span, personal feelings,

and so on. Rather than controlling for these factors as would be done with a traditional lab-based study, *“the researcher has to make sense of data in the wild, where there are many factors and inter-dependencies at play that might be causing the observed effect"* [209, p.59]. Here, the interaction with the prototype, as well as the internal and external factors from the user and the environment are taken as part of a holistic view of the user’s interaction.

Research in the wild is facilitated by the flexibility of digital technology as a medium for prototyping, as the state of the art digital technologies allow for functional, robust, working prototypes to be developed rapidly and enabling the research team to *“embed a whole range of innovative technologies in real-world environments in ways that were unimaginable a decade ago."* [209, p.58]. This fits with the rest of my methodological approach, as I am interested not only in how users interact with the agonistic interfaces (RQ2), but also how using agonistic interfaces changes and influences engagement and debate around socio-politic topics (RQ3), themselves topics which are embedded widely throughout the users context. Therefore adopting a research in the wild approach throughout the studies allows me to understand both of these perspectives.

There is debate within the HCI community around “in the wild" research, as it can conceptually mean different things. Rooksby documents how the term “the wild" came to prominence in HCI [211, p.10], and deconstructs what different researchers in the HCI community conceptualise as “the wild". For the most part, the idea of “the wild" is anything but the laboratory. Rooksby goes on to argue that as we are performing social analysis, doing so in a “natural" setting is the most pragmatic way to do this, but this does not completely preclude a laboratory setting.

Moreover, there are also discussions about the suitability of the term “wild" when being applied to participants lives [233]. It is argued that “in the wild" was conceptualised in the Western context, and thus carries an Othering connotation [67], and therefore we should, as a community:

> *“start a conversation around the terminologies we use to describe our research approaches and contexts. We consider it an ethical imperative to be conscious of the words we use to describe people and places"* [224, p.182].

Mindful of this, as my research encompasses issues of Othering and stigmatisation, and out of respect for my participants, for the remainder of this thesis I will refer to my work as deployed “in participants homes/lives" or “deployed in situ".

The Screenr and Spkr studies (Chapters 5 and 6 respectively) were conducted by deploying them in situ, where prototypes were given to participants for one month, with minimal engagement with myself or the rest of the research team, during that time. The Moral Compass and Spotting Guide studies (Chapters 4 and 5 respectively) employed some elements of this, with homework activities performed in between workshops, in their own homes with their own devices. A full discussion of these workshops is provided in section 3.2.1.

A methodological strength of deploying in situ comes from picking apart the multiple factors that are evident during deployment with participants. In doing so, we can understand the subtle interplay of factors between the designed prototype, the topics being discussed, and the social and political factors at play in each domain.

3.1.4 Addressing Assumptions and Biases

In a critically-engaged work such as this, it is important to acknowledge existing assumptions and biases that are prevalent in existing practices HCI and the design of digital technology, as well as those that I bring to the work myself.

Acknowledging, and subsequently managing assumptions and biases throughout the research process is also seen as an integral part of interpretative qualitative analysis, as noted by Braun & Clarke, where such methods “[do] *not treat this subjectivity as bias to be eliminated from research, but tends to involve contextualised analysis, which takes this into account"* [37, p.21]. Within the context of HCI, Vandenberghe et al. [241] make the point that when involved in practices such as CTP, we as authors should make obvious the assumptions that ground our work. By way of example they discuss Linehan et al.’s *Games Against Health* paper [144], which explores the underlying assumptions within ‘games for health’ research - particularly the *“unacknowledged conception of a player as a deficient or broken entity in want of repair. Rather than understanding the complex life worlds of players, scientific interest usually revolves around how well an unchallenged type of behaviour change has been facilitated through a certain game design method"* [144, p.2]. This example is

used by Vandenberghe et al. to demonstrate that the assumptions underpinning a design exercise have a profound impact on the object being designed, and the user outcomes.

One key assumption behind this thesis regards criticality, specifically introducing or supporting criticality with users, on the assumption such criticality is beneficial. Broadly, this assumption is supported by existing work around diversity of political opinions and the various conceptions of democracy and how one should involve themselves in it (as discussed in Chapter 2). More specifically, in the domains of inquiry of this thesis there is work that supports the assumption that increasing criticality is of benefit to people. Concerning socio-politicised social media, Brooker et al. [41] and Doughty et al. [75] demonstrated how a social media stream can become inundated with a single viewpoint, drowning out diverse viewpoints. Similarly, studies of reality TV viewers has shown those who enjoy watching such TV are less critical in their viewing [217], which is problematic given the strong socio-political messages contained in such programmes. The filter bubble and echo chamber effects have been highlighted as problems for those who engage with news and opinion, necessitating increased criticality and awareness of diverse viewpoints and opinions. Based on this assumption, I have however been cognisant when designing the prototypes engaging with criticality. In Moral Compass, Spotting Guide, and Screenr, I ensured there were times where users could *lean back*, and not engage with the prototypes, such as during advert breaks. In designing Spkr I allowed participants to nominate when the device would be active, to ensure they had time away from the device to relax in their home.

3.2 Data Collection & Analysis

The majority of data collection and analysis work conducted has been informed by qualitative methods, with a small amount of quantitative data used throughout to support the interpretation and understanding of how users have used the agonistic interface prototypes.

3.2.1 Data Collection

The primary source of qualitative data was talking to participants, via semi-structured interviews, including focus groups, and workshops. Interviews and focus groups engage participants in a conversation broadly focused around a specific topic, and I chose to follow a semi-structured format in all of these, as I wanted to explore both specific issues with respect to the use of digital prototype, and more broad discussion around the relevant socio-political issues. Whilst I wanted to cover roughly the same topics with each participant, I also wanted to probe more deeply around particular areas that a participant felt strongly about or wished to talk more about, and as such a semi-structured format for the interviews and focus groups facilitated that [198, p.269-271].

In Chapters 4 and 5 I used focus groups at various points in the studies. Focus groups are a useful method for encouraging discussion, as they facilitate the participants to share *"diverse or sensitive issues"* [198, p.272]. Given the focus on socio-political issues, bringing participants into a supportive environment with others would encourage sharing and discussion between them, as *"participants are empowered and able to make comments in their own words, while being stimulated by thoughts and comments of others in the group"* [207, p.285]. Moreover, the focus throughout on diverse participant groups meant that focus groups would bring together a diversity of viewpoints within the conversation. However, it should be acknowledged that focus groups are not immune from critique, as there are complex social elements invoked during them, such as peer pressure [205], and dominating personalities taking over [207, p.285]. In order to manage this, I used the considerations highlighted by Sanders & Stappers for focus groups with a generative purpose, i.e. a focus group intended to produce ideas, designs, or share thoughts. They describe that composition of the group is important, and that participants should have some common ground on which they can all relate, either in their background, or through some shared experience [215]. As I was giving prototypes to participants, I was therefore providing them a shared experience that they could talk about together - something I noticed especially in the final focus groups of Screenr, where participants had not met each other before (see Section 5.3).

In order to *"immerse the participants in making observations and reflecting upon the experience domains"* [215, p.160] I used *sensitising* materials, which are designed to *"stimulate participants memories, and to provoke their observations"* [215, p.161]. I used these sensitising materials to structure the discussion in the focus groups, and provide tangible things to discuss. In three of the studies (Moral Compass from Chapter 4 and Spotting Guide and Screenr from Chapter 5) usage data collected throughout the study was used as printed sensitising materials. This allowed me to ask direct questions about features of the usage data, and gave the participants a point of reference around which to base their answers. Word cloud visualisation of usage data were shown to participants during Spotting Guide and Screenr to studies, to both remind the participants, and act as a prompt for the interviewer (myself) to ask about specific behaviours during the studies. During Moral Compass study, the coding of tweets was done on paper by participants, as a way to externalise, and discuss, the experience of trying to apply a moral coding to a tweet. Further details of this process can be found in Chapter 4.

During all focus groups, and some interviews, I was assisted by one to two research assistants and/or interns. We all used the same semi-structured interview protocol, and having multiple interviewers were present was helpful given the semi-structured nature of the focus groups and interviews, facilitating the probing of participants' thoughts around the topics being discussed. For clarity, a full breakdown of who assisted with each interview and focus group can be seen in relevant chapters for each study, Chapters 4, 5 and 6. All interview and focus group schedules can be seen in the appendices (Appendices B, D, F, H).

All interviews were audio recorded using hand-held voice recorders. All audio was transcribed shortly after interview, either by myself (Moral Compass and Spotting Guide) or a university approved professional audio transcription company (Screenr, Spkr).

Data Logging

Preece et al. highlight the potential disconnect between what users say during interviews and focus groups, and what they actually do [198, p.273], and they espouse the use of alternate data sources, such as data logging, as a way to explore this.

This data *triangulation strategy* [79, p.151] complements, and juxtaposes the interview and focus group data with the participant's usage of the systems, revealing any differences or points of interest between them. As such I instrumented data logging into the digital prototypes used by participants. Given the digital context of this work, the majority of data could be collected by the servers that supported the digital prototypes, or could be accessed securely in a remote way. Data logging was conducted remotely in the studies where direct physical access to the participant's device was not possible, such as when it was being used in their home. The precise details of the implementation of the data logging are discussed in Chapters 4, 5 and 6. Data logging had a pragmatic benefit to the running of the studies, as it was possible to see what participants were doing with the systems, and to ensure there were no problems, either technically or with their engagement with the system.

During the implementation and use of this logging data I was mindful of the ethical considerations when instrumenting remote, unobservable data collection. A discussion of how these ethical considerations were addressed can be seen in section 3.3.5 on page 70. The usage data was used in Chapter 5 to understand which aspects of the Spotting Guide and Screener users engaged, allowing their usage to be mapped to broader theoretical "types" of user in a social system.

3.2.2 Data Analysis

I analysed the data collected using interpretive qualitative methods, which I approached with a relativistic perspective. This is particularly driven by the complex nature of social and political topics, existing within a complex social environment. Interpretative qualitative methods are founded on the subjective interpretation of reality. Such an approach is useful in the context of this work because it has, as Elliot et al. describe, an *"emphasis on understanding phenomena in their own right (rather than from some outside perspective)"* [79, p.147]. Indeed, Braun & Clarke discuss how *"interpretative aims to go further than descriptive analysis, unpicking the accounts that are given, and asking questions like 'What's going on here?' and 'How can we make sense of these accounts'? It tries to gain a deeper understanding of the data that have been gathered, and often looks 'beneath the surface' of the data"* [37, p.174]. My research agenda is exploring a new and emergent context,

and therefore interpretative methods allow me to understand how agonistic interfaces and socio-political issues can be designed for, and how they can be integrated within digital systems. Taking an interpretative approach to the data analysis is also advantageous to the exploratory nature of the work, as it can *"help uncover interesting and relevant research questions and issues for follow-up research"* [21, p.105]. This has allowed me to use the findings of studies to iterate the design of digital prototypes throughout this work (as exemplified in Chapter 5).

Interpretative qualitative analysis is popular in HCI work exploring socio-political issues, and have been used to analyse a wide range of contexts. For example, Vlachokyriakos et al. [243] explored how a novel voting interface would motivate residents to engage with local politics, Hussain et al. studied the personal and social impact of online video games [122], and Doughty et al. analysed the sentiments expressed on Twitter around a socio-politically oriented TV programme [74]. This wide range of contexts, and the focus of study, demonstrates interpretative qualitative analysis is suitable for interpretting and probing thoughts and feelings of participants, but also situations and scenarios that a technology intervention may evoke or provoke.

It is important to note that subjectivity is an integral part of interpretative methods, as it is accounted for, and openly acknowledged, as explained by Willig:

> *"The difference between a methodical interpretation of a text such as an interview transcript, and the researcher's subjective view of it is that the former is based upon a systematic, cyclical process of critical reflection and challenge of the interpreter's own emerging interpretations whereas the latter is the product of the author's unmediated associations and reactions"* [250, p.156].

As part of my critical technical practice approach, I have acknowledged my own assumptions and biases when approaching this work, (in Section 3.1.4 presented previously), and based on this have adopted analysis methods that recognise these biases, and as such own my perspective, by describing my theoretical orientation and my research interests (my values, interests, assumptions, and so forth), as described by Elliott et al. [79].

Thematic Analysis

A commonly used qualitative analysis method is *thematic analysis*, which aims to capture patterns within the data, that are subsequently clustered into a set of themes. Thematic analysis is commonly used in the HCI discipline, as well as other social science oriented disciplines. First developed by Gerald Holton, and popularised by Braun & Clarke [37, p.174-178], thematic analysis presents a method for *"identifying themes and patterns of meaning across a dataset in relation to a research question"* [37, p.175]. Braun & Clarke describe that there are two main approaches to thematic analysis, one *inductive* and one *deductive*. An inductive approach grounds the codes and themes resulting from the analysis in the data itself, and does not impose any pre-existing structure onto the data. This is contrasted with a deductive approach, where pre-existing theories are applied to the data, as a means to understand or view the data. I adopt an inductive approach to my analysis, which Braun & Clarke note is the most widely used of the two for data analysis [37, p.175]. This decision is also driven by the interdisciplinary nature of this work, as well as the context of socio-political issues. Inductive thematic analysis accommodates the subjectivity of the researcher, by grounding the analysis in the data, with the data providing the *"bedrock for identifying meaning and interpreting the data"* [230, p.22], and also positions the researcher as having a *"minimalist agenda beforehand* [...] *You observe life unfolding before you and construct meanings as they happen and later during your private reflections and writing"* [214, p.26].

Furthermore, thematic analysis is flexible as a method: *"it* only *provides a method for data* analysis; *it does not prescribe methods of data collection, theoretical positions, epistemological or ontological frameworks. It really is 'just a method'"* [37, p.178]. I wanted to ground my analysis in the thoughts and feelings of individual participants, who had experienced the socio-political issues being addressed, expressed their own viewpoints, and had also engaged with the digital prototypes. The flexibility of thematic analysis allowed me to analyse the variety of data generated from interviewing participants and their usage of the prototypes, to answer my research questions.

I did not conduct my analysis alone, and I worked with two other researchers (Gavin Wood and Scarlett Rowland) based in my research lab, who acted as sec-

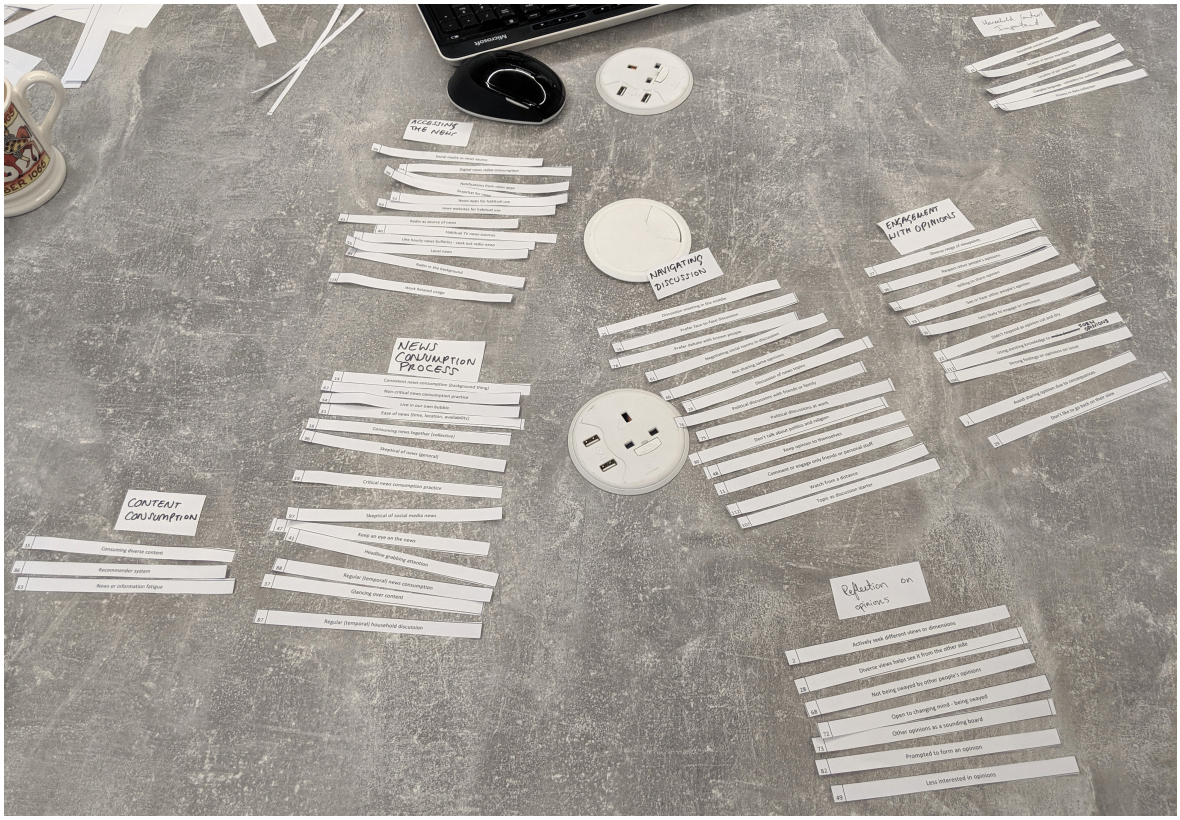

Figure 3.1: Clustering process for Spkr data, showing an initial sorting into seven themes.

ond, and where applicable, third, analysts. I wanted to involve other researchers in the analysis process given the socio-political context of the work. While it is common that interpretative thematic analysis is conducted by a single researcher, engaging multiple researchers in the analysis process allows the development of a diverse range of codes, precipitated by the diverse viewpoints that multiple researchers bring to the analysis [38, p.11]. Personal experience around social and political issues can be vastly different, therefore involving other researchers adds varied experiences and perspectives on the data, and helping to challenge my own assumptions and biases. Indeed, this resonates with Elliott et al.'s recommendations of good practice for qualitative researchers, to *"provide one or more credibility checks"* during analysis, which in this case I am doing by soliciting and comparing different perspectives on the data [79, p.157].

I used NVivo for all thematic analysis, as it provides specific functionality for creating codebooks, and annotating transcripts with those codes. It also provides reports that summarise the content of codes, and each coder could use NVivo to view the transcripts and use the same codebook as required. Specific details of who assisted me during the analysis, and their role in the analysis, can be seen in Chapters 4, 5 and 6. I took a complete coding approach, where I aimed to identify *"anything and everything of interest or relevance to answering* [my] *research question"* [37, p.206], thus remaining open to anything participants talked about.

Developing Codebooks Once all interview and focus group data had been transcribed, I commenced the coding process. My approach has been heavily informed by Braun & Clarke, who provide a comprehensive breakdown of creating a codebook and performing the clustering of codes into themes [37, p.207-274]. Drawing on these best practice guidelines, the coding process I followed for each study was broadly similar, and consisted of the following:

1. I familiarised myself with the data by reading it through. I also requested the other coder(s) read through the data also.

2. I created an initial codebook from an initial pass of the data. Depending on the size of the data this was either the whole data set, or a sample drawn from all the participants. I used an iterative process to generate the codebook, which used a mixture of data-derived codes (for example, using the exact words of the participant) and researcher-derived codes (informed by the theoretical knowledge of the researcher) [37, p.207], and as my initial coding pass continued I would merge, rename and split apart codes are required.

3. This initial codebook was then discussed with the other researcher(s). We discussed each code, whereby I gave a description of what the code encapsulated, and showed some examples of text that I had coded. If disagreements arose between the group we discussed them until reaching a conclusion about what to do with the code(s). I made notes based on the discussion, such as to rename a code, split them apart, or create new ones. After the discussion,

I adjusted the codebook as agreed, and I asked them to conduct a short initial coding (approximately 1 hour of coding time) with the codebook to see if they still agreed with the coding. During this period we would clarify and discuss anything that arose. At the end of this coding session, we discussed any changes arising, resulting in the agreed codebook.

4. I used the agreed codebook to perform a full coding of the data set. Once this was complete, I reconvened with the other researcher(s) to discuss any changes I made to the codebook (if any). We also discussed the content of the codes after the full coding. Once we were all happy, we moved onto the clustering stage.

Clustering Codes into Themes In order to cluster the resulting codebook into a set of themes, I met with the other researcher(s) in a meeting space. Working with other researchers whilst clustering was particularly useful as it provided multiple perspectives on patterns within the data, which could be articulated through the discussion. The aim of physically meeting together allowed us to see all of the codes in one place, as well as quickly and easily move them around. Figure 3.1 demonstrates this clustering process for the Spkr study (Chapter 6). Broadly, the process for clustering the codes into themes was as follows for each study:

1. I committed the final codebook to paper, either on sticky notes, or printed onto individual pieces of paper (see Figure 3.1 for the latter). The codes were laid out onto a large surface so everyone could see. Before we began clustering, we had a brief discussion to remind ourselves of the research questions/aims of the study.

2. We worked together to propose groupings for codes. This took the form of grouping only a handful of codes, through to large multiple groupings. We discussed the grouping, and why it should exist, and any tensions with existing themes. Theme names were collectively derived to encapsulate the codes within them. Throughout we discussed how the theme structure represented the data, and represented the research questions/aims for the study. In some instances we would discard all existing themes and start again, if we were

not happy with the theme structure. Themes could themselves be grouped together if relevant, creating first order and second order themes.

Further Iteration Following the clustering into themes, I wrote up the results using the thematic structure derived in the meeting. In doing so I also enacted minor tweaks such as renaming a theme to better encapsulate what participants were saying. I was mindful to maintain a narrative flow throughout this process, in order to *"tell a richer, more detailed story about* [the] *data"* [37, p.260].

3.2.3 Critical Reflections

As part of critical technical practice the critical reflections contained within this thesis, predominantly at the beginning and end of Chapters 4, 5 and 6, as well as in the Discussion (Chapter 7) and Conclusion and Future Work (Chapter 8) are considered part of the data. Although this data will not be analysed in the same way as the qualitative and usage data as detailed on the previous pages, it is important to acknowledge them as data that has informed my decisions, and that I will draw on them throughout.

3.3 Ethical Considerations

I will now discuss the ethical considerations that are presented by my methodology and research context. I will describe the process of informed consent, minimising harm to participants and other ethical considerations when working with human participants. I also approach ethics as *practice*, as proposed by Brown et al., whereby I consider not only the mandatory, legally-framed, ethical considerations, but also draw upon best practice and prior examples of the moral and ethical considerations being taken by the research community [42]. This is part of an *ethical responsive* process, as advocated by Durrant & Kirk [77], where I exercise reflexivity on my ethical position in designing and deploying the agonistic interfaces, and encouraging participants to enter into debates. In this section I will pay particular attention to the usage of *unseen* data logging through the prototypes, concerns around socio-political topics and emotional responses, and the usage of social media data.

Throughout, my work has been guided by the ACM Code of Ethics and Professional Conduct [1], and ethics-focused work in HCI. All of the studies in Chapters 4, 5, and 6 were submitted to, and approved by, the Ministry of Defence (MoD) Research Ethics Committee (MoDREC) as a condition of the funding of this work. Please see the declaration at the beginning of this document for full details.

All of the work comprising this thesis involved human participants. A core tenet when working with participants is to maximise the benefit, if any, to participants, while simultaneously minimising any risks to the participants. This is encapsulated by point 1.1 of the ACM: *"1.1 Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing."* To do this, in the following sections, I discuss the process for informing participants of the research, and maintaining their right to withdraw. Moreover I detail how they were compensated for their time given to the research studies, and how I protected their data. The ACM Code of Ethics also highlights another key tenet: *"1.2 Avoid harm"* [1], which can be interpretted to cover physical harm and emotional harm. With respect to physical harm, where participants were required to attend workshops, interviews, and focus groups, these were conducted in safe, risk assessed places. Any prototypes that were given to participants were electrically safe (physical hardware as in Chapter 6), and software was containerised so as not to interfere or cause harm with a participants device (such as their smartphone). I present a discussion of the potential for emotional harm and how this was mitigated in Section 3.3.7 on page 74.

3.3.1 Informed Consent

In order for participants to be fully informed about the research, I produced an extensive *participant information sheet*, written in lay language, which explained all details of the study. As part of the MoDREC process, I used the MoD template for participant information sheets, which can be seen in Appendices A, C, E and G. As a result, the information sheets were robust, in that they provided answers to a broad range of questions in clear language, e.g. "What are the possible disadvantages and risks of taking part?" Writing in lay language also helps to ensure participants, who are unfamiliar with research, understand what will be involved in the study. This aligns with ACM Code of Ethics point 1.3 - *"Be honest and trustworthy"* [1].

Once a participant had expressed interest in the study, as part of MoDREC standard procedure, participant information sheets were sent to participants, and they were requested to read them and consider the research for a minimum of 24 hours, and to respond via email if they were happy to proceed, or if they had any questions. At the first meeting with participants, I verbally reiterated the purpose of the research and asked them if they had any questions.

3.3.2 Right to Withdraw

Participants were reminded, in the information sheet, as well as at the beginning of the session (interview/focus group) that they had the right to withdraw from the study at any time, without consequence, and without giving a reason. Participants were also informed they could withdraw up to 30 days following the conclusion of the study, and in this case any data they had contributed to the study would also be removed. Participants were told that they would not be able to withdraw their data from a focus group discussion in the same way, as removing partial data from a group discussion would destroy the integrity of the focus group data. Therefore, participants were advised that should they wish to withdraw from the focus groups, there were to leave the room. I reiterated this at the beginning of each focus group, that they could withdraw at any point without giving any reason.

3.3.3 Data Protection

All data was stored on university machines, which are protected by username and password. In order to transport and synchronise data, cloud-based storage was used, provided by Dropbox on their Business tier. This service is password protected, uses two factor authentication and features a remote delete function. For any datasets infeasibly large for Dropbox (beyond 1GB file size), a portable hard drive was used. For security, the data was placed in password-protected compressed folders (.zip files) whilst on the hard drive, and was removed once it had been transported. When not in use, the hard drive was stored in a locked drawer on University premises. Given considerations for the *Data Protection Act (1988)* and the *General Data Protection Regulations* (GDPR) which superseded it in 2018, all data was

stored in anonymised state. Dropbox is covered under the EU-US Privacy Shield Framework, which provides adequate protection of personal data to the standard set out in the Data Protection Act (1988)[1], and GDPR[2]. For the purposes of GDPR (which was introduced whilst I was studying for this thesis), I was the data controller, as well as the data processor. The external transcription company, and other collaborators (as named on previous pages in this chapter) were also data processors for audio data and participant transcripts, respectively.

All data collected was anonymised, in order to preserve the privacy of the participants (as per point 1.6 *"Respect privacy"*, ACM Code of Ethics [1]). This involved anonymising all transcripts from interviews and focus groups, as well as removing identifying information, such as addresses, and telephone numbers for example. To further preserve anonymity, all audio recordings were destroyed following transcription. Section 3.3.7 contains a more detailed discussion of the anonymisation and protection processes implemented for social media data.

For each study, to aid organisation, and maintain the right to withdraw, a list of participants and their anonymisation number was created. This was stored in hard copy in a locked cupboard at Northumbria University, and was only accessible by myself. Thirty days after the conclusion of the study, this list was securely destroyed. After this it was not possible to identify a specific participant's anonymised data.

Data leaks are a risk when storing any quantity of participant data. To mitigate this, I enacted the anonymisation process mentioned, to ensure that all data was fully anonymised therefore in the event of a leak would not de-anonymise the participants. To mitigate the risk of a data leak, I ensured only myself and anyone relevant within the research team (see Chapters 4, 5 and 6 for specific details), had access to the data during a study. As noted previously, I stored all data with password protection. I gave specific consideration to social media data, ensuring tweets reworded and stored without their user ID, full details can be seen in section 3.3.7.

[1] https://help.dropbox.com/accounts-billing/security/data-transfers-europe-us
[2] https://www.dropbox.com/security/GDPR

3.3.4 Participant Compensation

Participants were compensated for their time volunteered to each study, according to how much time was required of them to complete the study. I gave careful consideration to the compensation rates for the in situ deployment studies (Screenr and Spkr), as they required significant involvement by the participants over the course of a month. To do this I calculated the approximate time required of participants per day or per week (as applicable), and combined this with time required for attending interviews/focus groups. The specific compensation details for each study can be seen in chapters 4, 5 and 6. Compensation was paid pro-rata, and should a participant withdraw mid-study they would have received appropriate compensation. In the event all participants completed the studies, and received the full compensation payment.

3.3.5 Data Logging

I implemented data logging into each of the digital prototypes, as a means to collect data about the participants' usage experience over the duration of each study. This data logging would not be obvious to the user of the digital prototypes, as the data logging was conducted automatically by the software supporting the prototypes. As outlined previously the prototypes were deployed to participants for use in the participants' daily lives as part of an in situ study. Automatically logging usage data throughout a participants day, without any visibility that their actions were being logged, raised an important ethical consideration, and as Brown et al. note we as HCI researchers should be mindful of the power we have when developing technology that participants will use [42].

Here a tension exists between preserving the participants' privacy by not collecting excessive data, whilst collecting sufficient usage data to understand how the prototypes were being used. I used multiple techniques in order to relieve this tension through the design of the prototypes and the study:

- During the implementation of each prototype I ensured the digital systems were separated completely (the principle of containerisation) at the technical

level, from anything related to the participant. In this way, it would be technically impossible for the data logging to log any data that was not related to use of the prototype. Here I drew on Crowcroft et al. [61], who advocate *privacy by design* for data logging, limiting the data collected from users to only what is necessary. This resonates with point 1.6 from the ACM Code of Ethics, which says *"only a minimum amount of personal information should be collected in a system"* [1].

For the digital prototypes that would be used on a participant's device (Chapters 4 and 5), the apps were developed as web apps, which would not have access to any data stored on the participant's smartphone or laptop, and could only log the participant's interaction with the web app directly. For the Spkr prototype that utilised hardware and software provided by myself (Chapter 6), the data logging was performed through the Amazon Echo voice recognition software, and collected audio from any user. To preserve participant privacy, this was handled by omitting any data that was not explicitly addressed to the research team. A full discussion of how this was handled can be seen in Section 6.2.2 on page 150. Participants were informed of this situation prior to the start of the study.

- I ensured that participants were informed of the data collection process before they enrolled in each study. Informing the participants, and allowing them to make an informed choice about whether to participate in the studies aligns with point 1.3 of the ACM Code of Ethics (*"Be honest and trustworthy"* [1]). At the commencement of each study, the participants were shown the information sheet, which detailed what, and how, data was to be collected about them, something which I reiterated verbally at their induction to the study, therefore making an explicit agreement around the data collection [198]. Given the considerations around the voice recognition used in the Spkr data logging, participants were told what the voice recognition would record, and of that data what would be used for the study, and what would be ignored (see Section 6.2.2 on page 150 for a full discussion of this).

3.3.6 Socio-Political Diversity and Viewpoints

Concerning the participants and the socio-political focus of this work, there are two ethical issues that need to be addressed: handling participants' possible emotional reactions to socio-political topics; and ensuring diversity of socio-political viewpoints within the participant sample.

Reactions to topics Indeed, by the design of these technology interventions, agonistic discussion filled with divergent and respectful conflict is actually encouraged and facilitated. However, given the ability for users to express themselves openly, both through the technology interventions, and in focus groups, it was important to anticipate that participants may become upset. To understand this, I have drawn on Saldaña's *Thinking Ethically* to orient myself to the potential emotional issues that may arise [214, p.79-92]. Whilst the debate takes place through the agonistic interfaces I have designed, engagement with these type of debates and topics is not unusual in the daily life of participants, and Brown et al. argue that *"making someone mildly uncomfortable, slightly irritated or momentarily inconvenienced should not constitute a barrier for research"* [42, p.9] given it is a feature of daily life. To this end, I integrated a number of steps into the methodology in order to identify and mitigate the risk of participants becoming upset:

- At the beginning of each focus groups and interview, I reminded participants that socio-political topics would be discussed and they might find them upsetting. Should this be the case they can withdraw from the session without giving any reason. In the case of focus groups they were told to leave the room to stop their participation. If multiple participants became upset during the focus group, I had a procedure in place to halt the focus group immediately, and institute a short break, after which time I could talk to the upset participants to see if they wished to continue. If there were other facilitating researchers (see each study chapter for full details of this), there were briefed about the above measures.
- During the studies, when participants were using the prototypes, I utilised the data logging to monitor participant usage, contacting a participant if they were

not using the prototype. This was done to ensure they had not become upset and stopped using the prototype. During the Screenr study (Chapter 5), I used the data logging to monitor discussion between participants to ensure nobody was becoming upset during the study. If this was the case, I would close the discussion using the administration interface. All participants were made aware of this at the beginning of the study.

I also provided printed contact details for myself, and for superior research staff, so the participant could contact them if they became upset during the study. I included my project principle investigator, and a researcher unrelated to my project (the Faculty Head of Department) should they wish to speak to someone outside the study.

Given the focus on agonistic debate between participants, it was important to define the boundary between agonistic debate, encouraged by the study, and an argument between participants. Agonism holds disagreement as central to a positive debate, i.e. people will disagree during debate, but this is ok. Further to this, agonism maintains that everyone in the debate is considered an adversary, rather than an enemy. As such, all participants in the debate are considered respectfully as equals with which to debate.

I defined an argument between participants as the point when the respectful nature of debate broke down, exemplified by *ad hominem* attacks to participants, or the use of abusive language. Of course, it is difficult to define an argument concretely, but I monitored participant debates using these cues. In the event, no participants became upset during the studies, nor were there any debates that were non-respectful.

Diversity of sample As the focus of this thesis is on critical reflection on real world socio-political topics, I was intent on recruiting participants who were drawn from diverse backgrounds, and who would naturally hold diverse perspectives and viewpoints.

The Moral Compass and Spotting Guide studies (see Chapters 4 and 5), used opportunity sampling, in the form of physical fliers, mailing list emails and word of

mouth to source participants. Participants were primarily drawn from the staff and students at Northumbria University, Newcastle University and University of Bath.

For the Screenr and Spkr (Chapters 5 and 6) I chose to use a recruitment agency, to increase the diversity of the participants recruited, and because these two studies required substantially more time from participants, as each took one month to complete, with almost daily usage. I used a local participant recruitment agency, who were able to source participants, drawn from different socio-economic backgrounds, ages, genders, and occupations, from the Newcastle upon Tyne area. Due to their diverse life circumstances, the participants had a wide ranging socio-political viewpoints. Striving to obtain a diverse range of participants throughout the studies is also part of the ACM Code of Ethics (1.4 *"Be fair and take action not to discriminate"* [1]), as using an agency goes some way towards involving those not familiar with research, or the topics being explored in this work.

Beyond the use of a recruitment agency, I did not recruit for diversity formally, i.e. I did not make inclusion/exclusion criteria based on diversity of viewpoint. This would be difficult because it would require me to know precisely which socio-political topic(s) I wanted different viewpoints on, and this could also work to exclude viewpoints or participants I had not anticipated. Therefore, in drawing from a range of professions, ages and socio-economic backgrounds throughout the studies, a naturally wide range of socio-political viewpoints were represented.

On reading Chapter 6, the reader may question my previous point that I have not formally assessed the diversity of participant socio-political views, as I performed political alignment of the participants. For clarity, the political alignment used in the Spkr study was applied to participants once they had been recruited to the study, and there was no selection for inclusion based on socio-political views.

3.3.7 Social Media Data

There are ethical issues that must be addressed when working with social media and social media data, most prominently the issue of anonymity for authors. Two of the studies in this work, Moral Compass (Chapter 4) and Spkr (Chapter 6) utilised data collected from Twitter as part of the prototypes. The British Psychological Society (BPS) have a clearly defined set of *Ethics Guidelines for Internet-Mediated*

Research that explain the ethical considerations, and suggest appropriate course of action for handling social media data [222]. I have used these guidelines throughout, and at the commencement of my PhD I used the 2014 guidelines, which have subsequently been updated in 2017. My work is also informed by the Association of Internet Researchers (AoIR), who published two versions of their ethics guidelines during my study: *Ethical decision-making and internet research: Version 2.0. Recommendations from the AoIR ethics working committee* [154] and *Internet Research: Ethical Guidelines 3.0* [89]. I have used all of these as supplement to the ACM Code of Ethics, as they provide more specificity around the unique issues encountered when conducting research with social media content. Drawing both of these guideline sets, I will discuss informed consent, anonymity, how to represent through publication, and mitigating the risk to users.

In order to guide this discussion of the ethics of social media data, I will discuss ways in which I have used it within the prototypes and within the studies, and subsequently explore each of the arising issues, and how I have addressed them.

Social media data within a prototype In the Moral Compass study (Chapter 4), I used tweets that were explicitly posted to the hashtags associated with TV programmes. These hashtags were defined by the TV programme makers themselves, in order to create a communication backchannel for their viewers (see Section 2.4.3 for a discussion of this). The study was configured to use live TV viewing, and also the live Twitter backchannel discussion taking place on the hashtag. As such, it was impossible to rephrase tweets due to the time sensitive nature of the study. However, to anonymise the Twitter users, their usernames and Twitter handles were removed, so only the tweet content was visible. The tweet content was only stored within the Moral Compass app, and was not visible to participants outside the confines of the study.

In the Spkr study (Chapter 6), tweets were used as the content for the Spkr device, which read out tweets in the home. All usernames were removed from the tweets to anonymise the authors. As good practice, all tweets were rephrased to maintain their meaning. This stopped participants finding the original tweet author. Participants were the only people who heard the content, as I did not perform analysis of the tweet content directly, and did not intend to publish them.

I will now discuss in detail the way I addressed the ethical issues arising from these two studies.

Informed consent online The technical systems that provide data from Twitter, the API (application programming interface), allow for the collection of publicly shared tweets from any user on Twitter. This is provided by Twitter as part of their terms of service, but it is not obvious to most Twitter users, and thus it is imperative to discuss consent.

Linking back to the discussion of social media as an online public (see Chapter 2), the consensus around informed consent and the use of publicly accessible social media data is that informed consent is not required. This applies to social media platforms where the initial expectation of the author was to share their tweets with the public domain, to be seen by anyone. Twitter is an example of such a platform. The BPS define this: *"Valid consent should be obtained where it cannot be reasonably argued that online data can be considered 'in the public domain'"* [222, p.9]. Therefore, using tweets that have been publicly shared does not require informed consent from the tweet authors. The AoIR also reinforce this: *"the greater the acknowledged publicity of the venue, the less obligation there may be to protect individual privacy, confidentiality, right to informed consent, etc"* [154, p.5]

Both note this should not be read as a *carte blanche* to use social media data within research. It is important to consider the authors' original expectation when they produced the content, highlighted by both AoIR and BPS:

> *"Also, participants may not be fully aware of the degree to which their discussion group posts are already available to public scrutiny, so making this clear in valid consent information may be appropriate."* [222, p.11]
>
> *"Do participants in this environment assume/believe that their communication is private?"* [154, p.7]

I only used tweets from Twitter, which have always been publicly visible. It is possible that an account can be set to protected[3], whereby only the account's followers are able to see the tweet. In this event, the tweets are not considered public domain,

[3]See: https://help.twitter.com/en/safety-and-security/how-to-make-twitter-private-and-public

and they are also unavailable via Twitter's API. As such, all of the data used consists of public domain tweets.

Users of Twitter may also delete their tweets, or their entire account, and thus all of their associated tweets. As the users will not know they are part of a research study, it is important to handle how their data might be withdraw, or suitably anonymised. I have addressed this possibility in the following section.

Anonymity & pseudonymity Anonymisation of social media data follows the same principles as anonymisation of any other research data (e.g. point 1.6 *"Respect privacy"*, ACM Code of Ethics [1]). Unlike traditional data collection *"situations where data are collected in* [Internet-Mediated Research] *with no potentially identifying information attached are not common."* [222, p.10] Therefore, I examined the format of all data collected to understand the identifying features. Then, once collected I ensured that all of the data collected was suitably anonymised. In the case of public tweets from Twitter, this consisted of anonymising the author, and rephrasing the tweet text.

Twitter contains two mentions of the author per tweet, this is both a screen name (e.g. Tom Feltwell), and a unique user handle (@tfeltwell). To anonymise all of the data collected, the screen name and user handler were replaced with participant numbers (e.g. P1, @P1).

It is also important to rephrase the tweet content, as it is possible to trace a tweet to its source using a search engine. This process is recommended by the BPS: *"for example, researchers may consider paraphrasing any verbatim quotes so as to reduce the risk of these being traced to source, and participants identified"* [222, p.14]. In line with this, all tweets were rephrased, so as to keep their original meaning, but not de-anonymise their author through a search engine.

Use in publication For publication, one strong recommendation of best practice made by the BPS and AoIR is to preserve the anonymity of social media authors in publications. To this end, I ensured that where applicable there would be no Twitter usernames featured in any written or printed documents. Participants did discuss Twitter accounts explicitly in some focus groups and interviews, and were

subsequently transcribed and used in published papers. In all of these instances the accounts were already anonymised.

Risk to Twitter users Throughout the consideration of ethical issues around social media data, a core principle was to protect the Twitter users from risk, by evaluating the risks in the collection and use of their data. Two key risks were identified:

1. Data leaks are a risk when storing a quantity of participant data. To mitigate this, I enacted the anonymisation and data storage processes outlined in Section 3.3.3 on page 69 and on the previous page.

2. A more remote risk was posed by the participants of the studies. During the studies participants would be interacting, via the prototypes, with the social media data. Given the socio-political nature of the studies and the tweets, a possible risk was for participants to become provoked by the content they were hearing, and want to identify the author. This was mitigated by presenting participants only with fully anonymised, rephrased content, to reduce this possibility.

3.3.8 Researcher Safety

As part of the in situ deployment studies (Chapters 5 and 6), it was necessary for a researcher to visit participants' homes, to conduct interviews, install devices etc. Therefore, it was important to consider the safety of the researchers visiting a participant's home. As the primary researcher, I made all home visits, and where possible I was accompanied by a fellow researcher. Exact details of who accompanied home visits for each study can be found in Chapters 5 and 6. Where this was not possible, I operated a buddy system with a senior member of staff, notifying them when I entered and left each home. In all instances, for safeguarding purposes, a senior member of staff had the full schedule, including names, addresses and timing of all home visits. I also carried my mobile phone at all times.

3.4 Summary

In this chapter I have presented my methodological approach responding to the research questions, which involved adopting a critical technical practice approach. This has led me to critically reflect on the assumptions and biases of HCI research with respect to choice of methods and desired outcomes from research studies. To actualise my critical technical practice, I have grounded my computer science technical practice in adversarial design, which also engages criticality in order to engage with *the political* - socio-political topics - and to design objects that make the user engage in *agonistic* debate. I have designed digital prototypes, that allow me to present users with real computational experiences, and I have adopted an in situ deployment approach to studying the technology, with participants using the digital prototypes in their real life context.

To study how the digital prototypes were used by participants I collected data using semi-structured interviews and focus groups, as well as usage data logged by the digital prototypes themselves. In my analysis, I have oriented my work towards interpretative qualitative analysis, which looks to understand the complex social environment that the data occupies. Thematic analysis is a popular method of interpretative analysis, and I chose to ground my analysis within the data itself, performing inductive thematic analysis on the interview and focus group transcripts. I worked with other researchers during the data analysis to draw on their diverse experiences of socio-political issues. The thematic analysis resulted in a set of overarching themes that describe the data.

In particular I see the discussion of ethical considerations in this space as useful to those wishing to work in this context. The interdisciplinary nature of working with socio-political content in the context of social media and digital technology is evident, as my discussion draws from multiple fields, which have various sets of ethical guidelines and best practices. Guidelines such as the ACM Code of Ethics are broad, generally unchanging, and are intended for the entire discipline of computer processionals. By contrast, research work around social media and Internet culture is rapidly evolving in response to new cultural and technical trends, as evidenced by the iteration of both the BPS and AoIR guidelines over the course of this

PhD. In the preceding chapter I have integrated all of these guidelines to present a comprehensive set of considerations.

Having established the context of this work, the relevant background literature, and the methodological approach being used, I will now move to discussing the domain of socio-political social media and the design of agonistic interfaces.

Chapter 4

Domain One: Socio-Political Social Media

The first domain of inquiry focuses on the consumption and engagement with social media content, specifically that concerns socio-political issues. Social and political issues are debated across social media around a variety of issues, such as elections [43], political protest [231] and cultural phenomena [105]. One topic of recurrent debate in the UK centres around state welfare provision [129].Some TV programmes specifically focus on state welfare, where they generate large discussions on social media. This genre, a sub-genre of reality TV, often dubbed 'poverty porn' [127], focuses on people who are supported by state welfare. There is often discussion on social media about the subjects of these programmes. Indeed, programme makers have grown savvy to these online discussions, with many overlaying an 'official' Twitter hashtag at key points in the programme to drive viewers to the same online discussion space [75]. The recent proliferation of this genre has led to concerns that prime-time media experiences are exacerbating misconceptions, and stifling critical debate, around major societal issues such as welfare reform and poverty. Previous work has demonstrated how the socio-political issues debated around these programmes are often lacking in diverse, critical, perspectives [41], and thus presents an opportunity to explore agonistic interfaces to improve this diversity.

In this chapter I present a study exploring an alternative representation of Twitter streams that uses 'morality' to represent the different views being raised. To this end, I designed a smartphone app, Moral Compass, that allowed users to explore this representation through a novel, compass-like, interface. Moral Compass was

deployed through a set of three workshops, where participants explored assessing the 'morality' of tweets, and were paired up to assess the the morals of each tweet within a live TV-programme related Twitter stream, and subsequently explore this representation live through the compass interface.

4.1 Existing Practices

As I introduced in Chapter 2, social media facilitates much socio-political discussion, which is achieved through a variety of means. A prominent feature is that of aggregation, with many examples of hashtag aggregating content leading to social movements to organise. This can be done consciously by an activist (e.g. in counter discourse activism as discussed by Feltwell et al. [84]), or happen organically by multiple users who use the same hashtag. Both situations have the same effect, creating one large online discussion space, which has been used for large scale societal social movements. For example the Gilet Jaunes protest movement (as discussed in Section 2.4.2 on page 36) is an example of an organically organised social movement, who have been documented for their use of social media as a means to discuss the underlying socio-political issues at the centre of their movement, as well as using hashtags to incite action.

Social media features such as hashtags are often used in complex ways by a variety of different actors (activists, users, the media, and so on), who are all able to shape the discussion around a hashtag. Discussions can sometimes spawn other related hashtags, as well as being associated with a temporal event, or a more broad cultural movement, and are often used in combination [124].

The hashtag, as introduced in Section 2.4.2, is a user-generated tag that is used to aggregate content together, e.g. all photos tagged with #dinner will appear when a user searches for #dinner, irrespective if the poster is a friend, etc.

Much of the utility of hashtags comes from their user-generated nature. Twitter only moderates individual posts that are deemed to break "The Twitter Rules"[1], but

[1] Twitter. `https://help.twitter.com/en/rules-and-policies/twitter-rules` Accessed 18 January 2020.

beyond this, they do not moderate discussion: *"as a policy, we do not mediate content or intervene in disputes between users."* [2]. Thus, provided a hashtag (and the associated post) are within Twitter's rules, any combination of hashtags, existing or new, can be posted, and Twitter will not intervene. As a result hashtags are utilised as a tool with which to exert influence or, in the terms of Castells, to enact communication power [51]. The malleability of Twitter hashtags thus allow other actors such as activists to purposefully target an existing hashtag to obtain control over it [111], or to connect their message with the existing message of the hashtag, known as frame bridging, as a means to disseminate their message to a larger audience [84].

As previously discussed, Brooker et al. [41], amongst others, have highlighted how the aggregation of all tweets around a hashtag can lead to overwhelming negative reactions dominating the feed, based on the quantity of tweets. Brooker et al. note that there *is* diversity within these streams, but they are in effect drowned out by the interface, which does not select for viewpoint diversity, but for volume.

Considering how such an algorithm might be reconfigured, it is fruitful to examine those who already leverage Twitter for their own devices. For the purposes of marketing, there is substantial knowledge and information regarding the way that such algorithms function. Indeed, an understanding of the way to leverage and manipulate the algorithms of social media platforms is integral to the success of marketing campaigns, and for those users considered *influencers* - users of social media who act as brand advocates for companies [83], who often demonstrate an understanding of how hashtags work in order to increase the number of followers they have, thus increasing their appeal as a brand advocate [3]. Over time Twitter has undergone a number of changes to their feed algorithm. When the work in this chapter was initially conducted (2016) Twitter had only recently introduced an algorithm to their feed (see footnote[4] for an explanation of these changes). Prior to 2016, all tweets were viewed by users in real time, without any algorithmic filtering. Twitter have reimplemented this feature, as a user interface option, from 2019. Twitter, as

[2]Twitter. About Offensive Content. `https://help.twitter.com/en/safety-and-security/offensive-tweets-and-content`. Accessed 20 January 2020.

[3]Influencer MarketingHub. `https://influencermarketinghub.com/the-ultimate-guide-to-using-instagram- hashtags-to-grow-your-followers/` Accessed 20 January 2020.

[4]Hootsuite. `https://blog.hootsuite.com/twitter-algorithm/` Accessed 20 January 2020.

with many other social media companies does not publicise the exact workings of their algorithms, in order to avoid manipulation of their algorithms to the detriment of the platform. Given this secretive nature, users therefore develop folk theories, that explain how the algorithms work from their perspective [66].

An example of a system that leverages social media hashtags is the now defunct service Thunderclap[5]. Thunderclap was a crowdspeaking platform designed to amplify a message, usually created by an activist, on social media. It utilised the Twitter API to post the message automatically, once the campaign had a certain number of supporters. The effect was that large numbers of users would post an identical message at the same time, and I previously studied how hashtags can be used in a Thunderclap campaign to flood an established hashtag, as a means of injecting a counter-discourse [84]. In this way a tool like Thunderclap can be used to overwhelm the Twitter algorithm.

It is important to say I do not intend to frame Twitter's policy choice not to moderate as näive. I believe it is most likely intentional, to encourage people to use Twitter as a space for discussion, without any input from them beyond enforcing their rules, in line with a libertarian perspective [229].

A further behaviour stemming from Twitter hashtags being unmoderated is the use of hashtags by TV producers to channel users towards a shared discussion space. This practice of displaying a hashtag on screen, or explicitly mentioning the hashtag on the programme is part of a *cross media strategy* to drive viewers to engage with the programme on social media [64]. One means by which public reaction to television might be measured is by analysing the nature of the social media backchannel [75] discussion to its broadcast. Live backchannel discussion of TV has been studied in many contexts, such as major sporting events [186], and also around politically-themed content such a televised debates [107]. The research has shown that whilst backchannel discussion might support some reflective and pluralistic dialogue [75], it may also lead to trolling and other forms of online abuse [74], which aligns with the findings of Brooker et al. presented previously who advocate the redesign of the Twitter interface to introduce more viewpoint diversity [41].

The extent to which the framing of TV content influences viewers' attitudes is a subject of long-standing scholarly debate [48, 125, 179]. However it is clear that

[5] `https://en.wikipedia.org/wiki/Thunderclap_(website)`

when combined with the nature of the Twitter feed, prioritising for most recent rather than viewpoint diversity, this creates a space for negative reactions, where Twitter users are not able to see a balanced discussion. In line with this idea, oriented towards reconfiguring news feeds, Baumer et al. describe a visualisation tool, Reflext, that allows users to explore socio-political issues [15]. It algorithmically derives how socio-political topics are framed across a variety of mainstream news outlets, allowing the user to see common word pairings or association, and reflect on this framing.

This section has demonstrated that there is a design opportunity presented by the Twitter interface, particularly when users are engaging with socio-politicsed hashtags. Creating a space for reflection on the hashtags contents, such as by revisualising the hashtag's constituent tweets using a different metric rather than how recently they were tweeted, is a promising design idea.

4.2 Moral Compass

Taking lessons from the preceding discussion, I designed and deployed a smartphone application, Moral Compass, aimed at encouraging the provision and creation of user-generated content around these socio-political social media streams, on the grounds that it promotes engagement and reflection by users. Moral Compass is premised on encouraging less passive and more active engagement with Twitter, utilising interactions such as 'tagging' and 'navigating', as opposed to merely passively reading, content. This, according to the research discussed previously, has the potential to support lasting engagement with the socio-political issues emerging from broadcast media. In the following sections, I describe the design and implementation of the Moral Compass, the study design, the findings of the study, and a discussion of the results.

Designing Moral Compass

The design and initial implementation of Moral Compass was conducted together with Tom Schofield, Senior Research Assistant at Newcastle University. We worked collaboratively to reach a design that allowed the re-configuration of a live Twitter

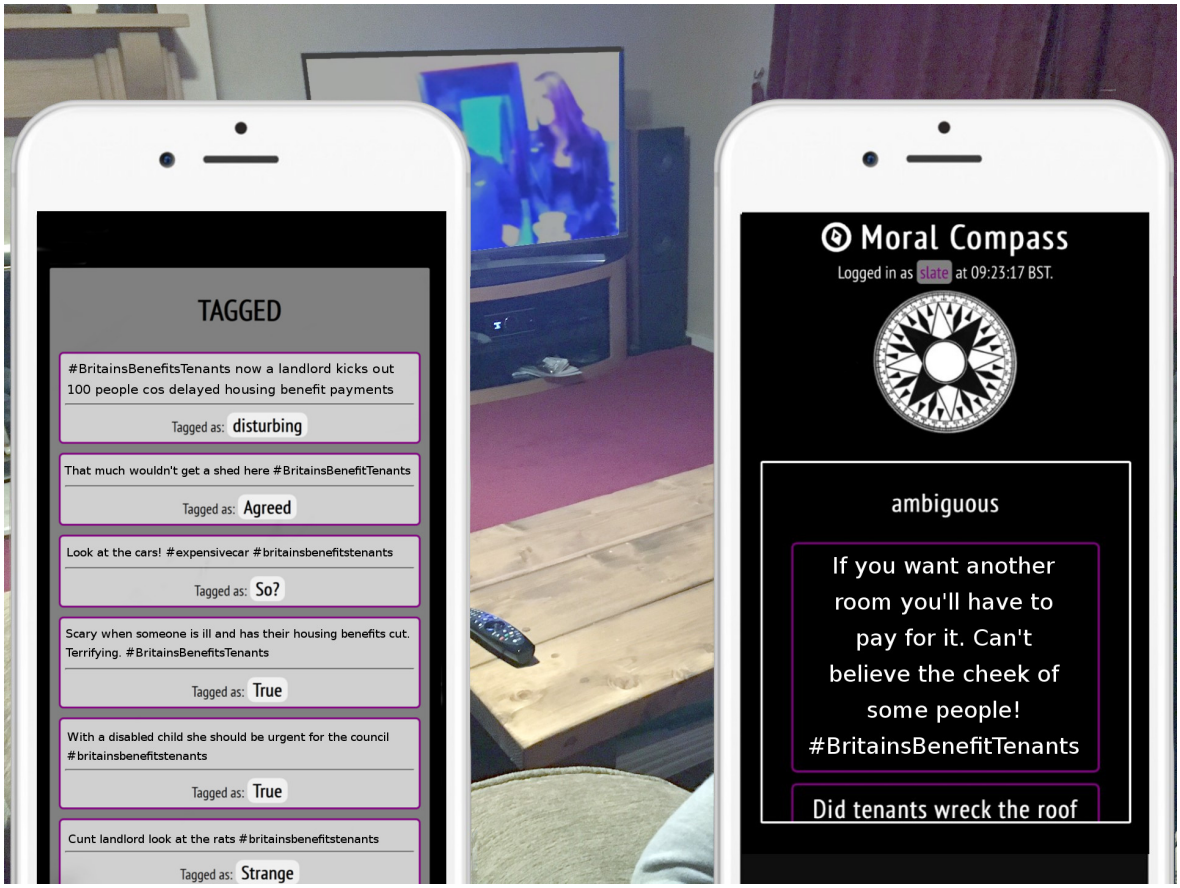

Figure 4.1: Mock-up of Moral Compass Tagger (left) and Moral Compass Explorer (right)

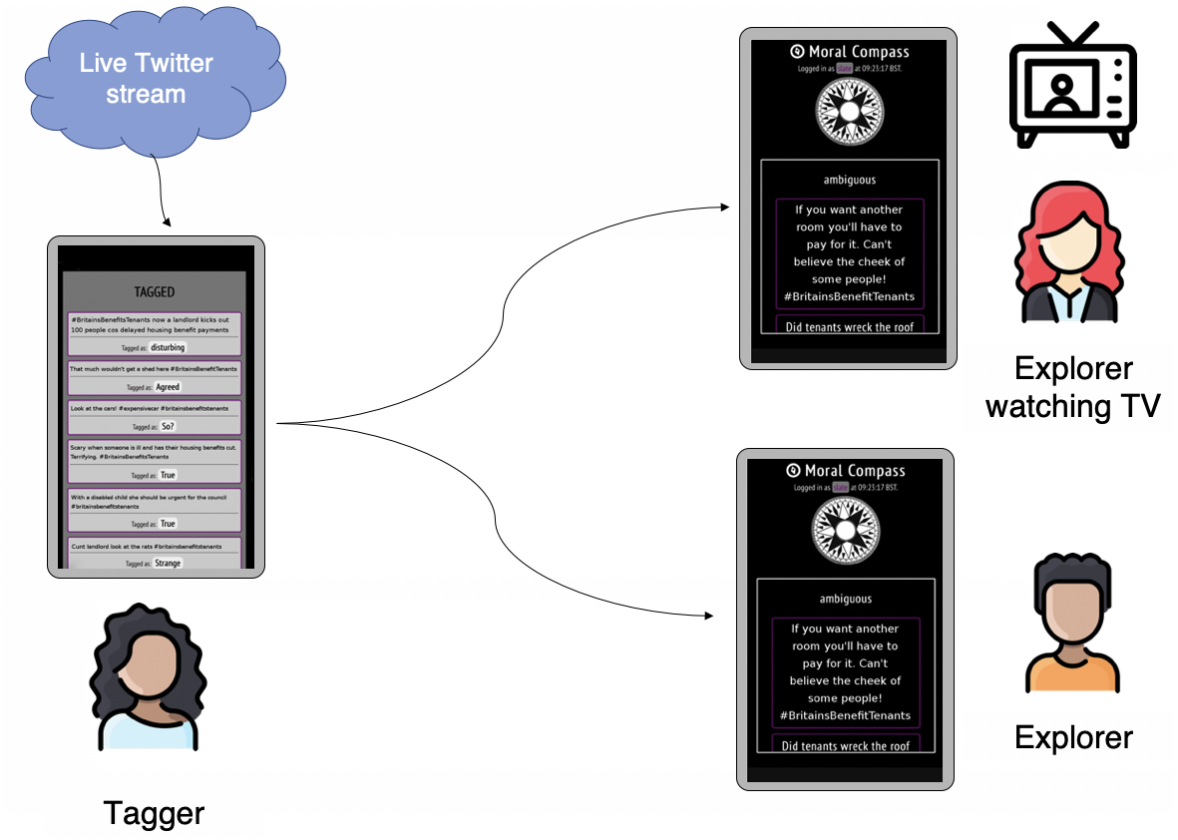

Figure 4.2: System functionality between Moral Compass Tagger (left) and Moral Compass Explorer (right). *Icons made by: Freepik, Vitaly Gorbachev via www.flaticon.com*

stream. This resulted in an initial prototype of Moral Compass. Tom built a concept interface for Moral Compass. After this, I took over the technical implementation and all future design decisions. The Moral Compass is a two-part app: Moral Compass Tagger for tagging live Twitter streams, and Moral Compass Explorer for visualising the tagged Twitter streams. Given that Twitter streams can display a plethora of political [41] and moral views [128] a user of Moral Compass Explorer watching a programme is able to see what the *morals* of the Twitter stream are, created by another user of Moral Compass Tagger. Of course, a *morality filter* does not exist as an interface, or data layer on Twitter, and thus Moral Compass Tagger generates this data by allowing users to tag individual tweets in a live, programme related Twitter feed with their interpretation of each tweets morality. Thus the Moral Compass Tag-

ger allows users to interpret Twitter streams, with Moral Compass Explorer allowing reflection on views of the tagger, and the views in the Twitter stream. Figure 4.2 shows how these two parts of the system work.

Implementation Unique usernames were used throughout the Moral Compass system, with all data stored against this username. Moral Compass Tagger displayed tweets from a selected Twitter stream, (e.g. a keyword or hashtag) predetermined by myself prior to the study. Tweets are displayed sequentially on the interface, and the user can click on each tweet to assign a tag. Tags are entirely supplied by the user through a free text input, and can be reused. Figure 4.1 shows the tagging interface. Users also had the option to ignore a tweet if they were not able to tag it, and it would be discarded. The list of tweets would be updated automatically as new tweets were posted on Twitter, and thus users were asked to check the app frequently and tag tweets at their own rate. Tagged tweets would then be available to the Moral Compass Explorer

The Moral Compass Explorer visualises the resultant morally-tagged Twitter stream onto points on a compass wheel, which a user is able to rotate. As they rotate the compass (with their finger or a mouse) the different tagged moralities appear, with their associated tweets (e.g. Figure 4.1 showing the “ambiguous" tag, Figure 4.3 showing “two sides" and “unfair"). For the purposes of the study I manually assigned one-tagger to one-explorer, but the app can also support one-tagger to many-explorers and many-taggers to one-explorer relationships.

Both parts of Moral Compass were developed as a web app, using HTML5 and JavaScript. STOMP and Apache ActiveMQ were used to supply real-time data from the Twitter Search API. Tagging data was stored centrally, which was subsequently printed to aid discussion in the workshops. To anonymise Twitter data, I followed the BPS Guidelines for Internet-mediated research [222] ensuring the use of only publically visible tweets associated with a hashtag of a TV programme, removing the identity of tweet authors and disabling tweeted hyperlinks as discussed in Chapter 3. Due to the app’s focus on the nuances of tweet content and speed of operation, I did not rephrase tweets seen by participants, however all examples in the figures and discussion presented here are anonymised.

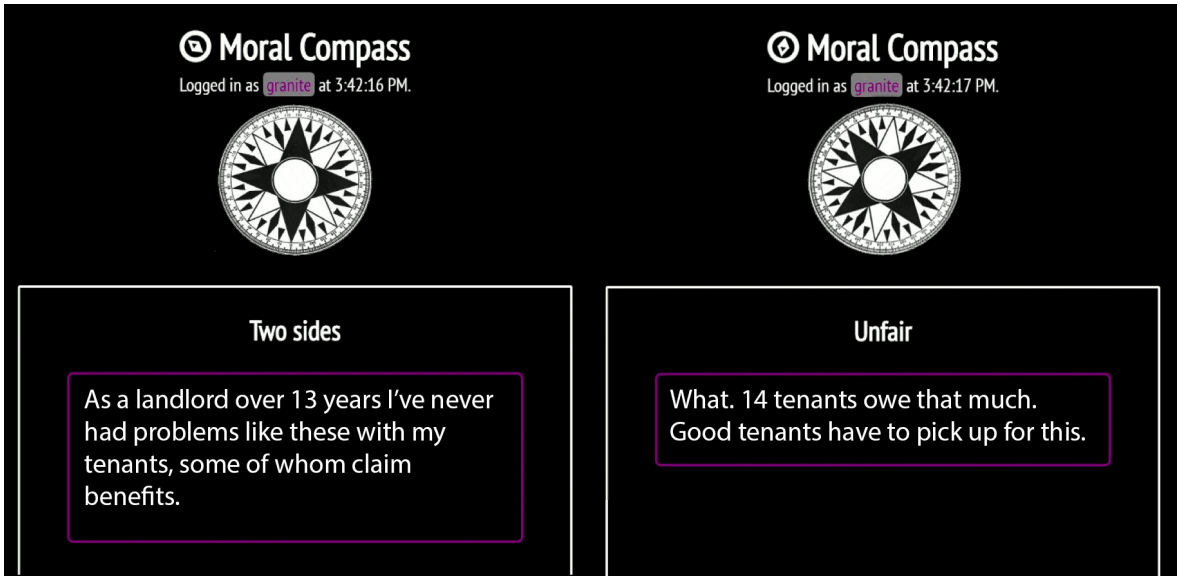

Figure 4.3: Moral Compass Explorer at two different compass points, showing tweets under the different moral codings.

Study Design

Participants were recruited using posters and email advertisement at the University of Bath, Bath, UK. To maintain discussion group sizes and to aid scheduling, participants were split into three separate groups. A total of 15 participants took part in the study, divided into three groups of five, four and six participants, totalling nine workshops. The majority of participants ranged between 18 and 50 years of age. The participants comprised 10 who identified as female and four who identified as male, and per group: three female/two male, three female/two male and four female/one male. All were familiar with, and used social media. Participants were compensated for their time with a £40 voucher. The study was conducted over a three-week period, with one workshop per week. Each workshop involved a discussion, short activities in relation to the app and “homework" between each workshop. Participant experiences and outcomes of their activities would then be discussed in the following workshop. Overall this formed three workshops and two homework activities per group. I was assisted in the workshops by Gavin Wood, Senior Research Assistant within my research group, and Phillip Brooker, Senior Research Assistant at University of Bath.

- **Workshop 1:** The context of the work was introduced, followed by a short discussion to explore the participants' knowledge of reality television and assumptions and experiences of othering and stigmatisation on social media.
- **Homework 1:** Participants were asked to use Moral Compass Tagger at a specific time where they would tag tweets from a live Twitter stream for approximately one hour, as a training and learning exercise.
- **Workshop 2:** Began by inviting participants to discuss their experiences from the homework task. Discussion then moved to the specific tweets they came across, broadening to their overall experience using the app.
- **Homework 2:** Participants were paired together, with one person tasked with making tags using Moral Compass Tagger, and the other using Moral Compass Explorer during a live broadcast of Britain's Benefits Tenants, a UK reality TV programme about those living on state welfare and their interactions with their landlords, focused on the worst examples of neglect or mistreatment of housing. Participant pairs were picked randomly considering those who had access to live TV (two of the participants did not have a TV set and/or license), with groups with odd numbers of participants having multiple Moral Compass Explorer users (as seen in Figure 4.2).
- **Workshop 3:** Started with a group discussion in their Tagger/Explorer pairs, followed by a wider discussion about the homework tasks and the overall reflections.

4.2.1 Data Analysis

During the workshops, all of the discussions were audio recorded and then subsequently transcribed. Inductive thematic analysis was conducted [39] on this data. Two other researchers (Gavin Wood and Scarlett Rowland) assisted as second and third coders. I produced an initial codebook based on an initial pass of the data, I then discussed this codebook with the other two researchers, until we all agreed on the codes. I then used this agreed codebook to recode the data. At this point, we collaboratively clustered the codes to create a thematic structure for the data.

4.2.2 Findings

Below I present the findings from the analysis, which are grouped around two overall themes of *Interpretation and Critique* and *Navigating the Compass*.

Theme 1: Interpretation and Critique

This theme covers participant's discussion around interpreting the content being presented to them in the Moral Compass Tagger, and on the process of making critique on tweets and its wider impact on the apps. This theme is divided into two sub-themes: *Interpreting Content* and *Critique and User Perspective*.

Sub-Theme 1: Interpreting Content Representing tweets in Moral Compass involved anonymisation by removing the author and disabling hyperlinks. Participants expressed that they struggled with interpreting tweets, noting there may be a deeper context: *"I thought that's not fair. I genuinely thought there's more to it."* [P1, G1]. This also manifested as a desire to follow up tweets to explore the issues being discussed by the tweeters: *"There are things I think I would really like to know, for example how the housing situation has improved in certain places"* [P6, G2] and *"Sometimes I want to look on the* [Twitter] *website to look at whole contents, but* [hyperlinks] *didn't work - it made me most confused"* [P14, G3]. Reflecting on the second screening nature of the Moral Compass Tagger, the activity placed an attentional demand on the participants: *"I was very disappointed in myself* [...] *I found that I couldn't multitask"* [P1, G1] and *"Sometimes I missed out like a couple of minutes, maybe, because I was trying to go through a lot of the tweets."* [P6, G2]. However, others noted they balanced this attentional demand by using different methods, such as a "catch all" tag or using the ignore button: *"Something like 'I love cannabis', I just put as 'irrelevant' because it didn't really say anything about the show"* [P6, G2] and *"I just ignored... some of them were so weird!"* [P9, G2].

It was clear that some participants were attempting to provide a balanced viewpoint of the Twitter stream through their tags. During Homework 2, one participant noted the tags used for the two-sided debate seen in the Twitter stream: *"I had 'landlord sympathy' but sympathy with tenants was just 'sympathy'"* [P13, G3]. When presented with their own tags in Workshop 3, generally participants noted they were

happy with what they had done: *“It’s funny to see the things that I’ve tagged them as but I do, I think I kind of stand by the things that I’ve said."* [P13, G3].

Sub-Theme 2: Critique and User Perspective The discussions around Moral Compass Tagger highlighted the presence of a tension between anonymised tweets and de-anonymised tweets. One participant described how anonymised tweets stopped prejudices forming about the tweeter: *“seeing them anonymously, I found it a lot easier to be impartial"* [P13, G3]. However, this was contradicted by others: *“I think I wouldn’t find much utility in that information without knowing* [...] *who they are."* [P10, G3]. This reflection around prejudice also lead to participants openly discussing their processes of self-censorship, or considering who might be viewing their tags: *“Actually it was me trying to police myself because two people were going to be reading it."* [P1, G1] and *“I had to tag it as ‘mean’ because if the tenants* [in the show] *... read this they would feel like they are being targeted"* [P6, G2].

One prominent discussion thread was the possibility of the user’s perspective and political viewpoint being reflected in the way they tag tweets: *“Somebody who’s quite right-wing or not in the same age bracket would have very different opinions and interpretations of what people were saying"* [P8, G2], and similarly, *“I would definitely be hesitant to read the rest of it, if I saw that somebody was tagging a really mean tweet as a funny one. I would think that the tagger is twisted"* [P6, G2]. Reflecting on a similar theme, another group suggested the problems that might be created when nominating a tagger. One participant sarcastically proposed, *“I want Stephen Fry to basically tag a load of things and make sure that I don’t see the things that he thinks are offensive"* [P13, G3]. In this vein, the same participant identified that the Moral Compass would be a useful tool for censoring specific viewpoints out of the Twitter stream: *“I was wondering during using it for the tagging thing about the potential for it to be used for censorship and I don’t really like the idea of that"* [P13, G3]. An interesting behaviour was noted in relation to the Moral Compass Explorer users, where they would initially try to understand the political viewpoint and perspective of the person performing their tagging. This process of trust building was explained by one participant: *“*[I] *probably looked more at the more controversial words like the stereotyping"* [P2, G1], to see if they would have tagged them the same as themselves, while others noted it took them time to become familiar with

their tagger: *“I think there’s a sort of period of time where your kind of understanding the consistency* [of the tagger]*"* [P13, G3].

Theme 2: Navigating the Compass

This theme focuses on the participants’ use of and reflection about using Moral Compass Explorer, and specifically about the compass interface. Participants felt the compass was useful for viewing a large number of tweets, due to tweets being shown by turning the compass wheel: *“it was quicker and easier to look at the compass than it was to look at the timelines"* [P2, G1] and *“it was less effortful, you can consume more, rather than, Twitter when you’re constantly scrolling up and down"* [P10, G3]. The ease of use was further drawn out when talking about the second screening experience: *“you could just almost flick through, look up, look down so it was quite easy to use."* [P2, G1] and *“I also liked that it wasn’t too distracting"* [P13, G3]. The compass interface was also used in playful ways, to find unexpected or new tags and tweets: *“Certainly playing a bit. Sometimes scrolling a bit further... maybe by taking a whole turn back and forth, just to see where I would end up"* [P7, G2] and *“I wasn’t purposely thinking ‘I’ll go back and look at...’ It was more going round and round and round"* [P11, G3]. Some participants noted that ad breaks in the programme provided them time to reflect on the tweets and tags around the compass, without distraction from the programme itself: *“I’d say as the programme progressed, I probably in the ad breaks paid more attention to it because of the... tweets that were being generated"* [P10, G3] and *“That’s why I liked the commercial breaks because it actually gave me a break to go back to the compass"* [P7, G2].

Significantly, participants described how they used the compass as a filter to look for tags of a specific kind: *“I wanted to look at the more positive comments so I tended to stop on some positive tags"* [P11, G3]. This sentiment was extended to second screening contexts where this would be useful: *“if I were watching, I don’t know ... Strictly Come Dancing tweets, or Great British Bake Off tweets, I wouldn’t want to see, in that context, racist tweets"* [P10, G3]. Subsequent to this, there was discussion about the qualities of different types of programme-related Twitter streams, and where debate fitted into it. Talking about cookery programmes: *“I just think that in some contexts, you don’t necessarily want to see that, because*

it's not a debate" [P11, G3], was contrasted with the stream for Britain's Benefits Tenants: *"something like tweets about this programme, feels like a debate and it's quite appropriate"* [P11, G3].

4.2.3 Discussion

In this study I set out to design and deploy an agonistic interface, via a smartphone app that facilitates and promotes more critical interpretation of social media streams. I explore how the findings, in tandem with previous literature, demonstrate explicit approaches that future platforms might use to encourage and facilitate critical reflection on social media streams.

Interfaces for Discovery and Filtering

At an interaction level, the compass interface of Moral Compass was found to be easy to use, and also allowed the serendipitous discovery of tweets & tags. The users toyed with the compass wheel, which led to a process of serendipitous discovery of tweets and tags some had not seen before. Future designs of second screening apps oriented towards reflection could scaffold this behaviour, e.g. encouraging serendipitous interactions [132]. It was also shown that participants used the compass wheel as a way to filter the tweets. By mapping tags to specific points on the compass, users were able to identify or avoid a tag they did not want to see, such as racist tweets. This process also extended to switching between types of content during the second screening process, moving between tags such as *humour* and *funny* to more critical and reflective tags such as *unfair* and *more to it*.

Balancing Attentional Pressure

Participants also reported feelings of pressure and harriedness in their tagging and exploring. Those who were watching a TV programme and interacting with the Moral Compass Explorer, expressed that using the app to view the coded tweets was intense, requiring them to rapidly switch their attention between the programme and the app. Using Moral Compass Explorer involved rotating and exploring the tags and tweets around the compass, which is an activity users can switch into and out of

as they desire, with advert breaks being noted as a natural moment to interact and reflect more deeply on the tags and tweets. Users of the Moral Compass Tagger described how they were able to tag tweets at a steady rate during the programme, with the *ignore* button helping them skip those that proved too difficult, reducing their attentional load. The design of both parts of Moral Compass demonstrates the effect different interactions have on guiding the type of reflection and discussion that occurs around a social media stream. This suggests that if a second screening app designed to provoke criticality and reflection was *purely* focused on criticality, it may not be as successful as one that allows users to move between critical, reflective content, and more humorous (though less reflective) parts.

Providing context

In re-visualising the Twitter stream in Moral Compass, and Moral Compass Tagger, participants felt that due to the anonymisation process and brevity, large amounts of context was lost. One of the affordances of the Twitter platform is that it allows users access to metadata (e.g. biographical information, recent tweet history, etc.) to contextualise and help situate the meaning of tweets. As participants in the workshops expressed, following up these details provides a potential strategy by which to glean extra information about a tweet and its author: e.g. their ideological standpoint, whether they troll, or who they follow. Without access to this extra contextual information, I surmise this may have contributed, in part, to the difficulty experienced by participants in tagging tweets. However, in line with “friction" as a design principle [135], it could be said that the lack of context required users to reflect on the meaning of tweets to a greater degree, as opposed to having those meanings easily explained by associated metadata. Hence, though providing some degree of context around information may be necessary, this should not be at the expense of providing opportunities to think about and critically evaluate and reflect on the information provided.

Support differing viewpoints

The unmoderated nature of the tagging activity was acknowledged by participants as allowing different opinions to be shared. Moral Compass was recognised as a

useful tool for filtering such content, which could therefore also be considered as a tool for censorship. These concerns are legitimate, as if only a single user performed the tagging, this may be the case. During the study I paired participants, with one tagging and one viewing the resulting tags. In a future iteration of this approach, multiple interpretations of the tagging, by multiple users, could be made visible to users of the Moral Compass Explorer, allowing them to explore multiple different compasses - representing different users' tagging of the same Twitter stream, and exposing them to a range of diverse viewpoints and perspectives.

4.3 Summary

This domain of inquiry has focused on socio-political social media streams. First I detailed the existing practices for the design of social media, principally how hashtags encourage the consolidation of discourse, and how certain external organisations (e.g. TV producers) are using this to their advantage. Based on this I discussed how online backchannels manifest, afforded by the design of, notably, Twitter. Motivated by this, I presented the design of a smartphone app, Moral Compass, which facilitated and promoted critical engagement with Twitter, by allowing users to encode their own interpretation of the morality of tweets. The findings demonstrate promising opportunities for the design of agonistic interfaces that provoke reflection and criticality towards socio-politically charged social media streams. Further, I have shown that summarising one's perception of tweets in a Twitter stream can be a useful way to encourage and facilitate more thought and reflection regarding the framing of social media stream content, and show how social tagging can evolve into a reflective process engaging multiple users.

Chapter 5

Domain Two - Second-Screening of Reality TV

As demonstrated in the previous chapter, reality TV is a prominent cultural site for socio-political issues to arise, and this is manifest not only in the associated social media streams, but also in the viewers of the programmes. This domain focuses on the TV media itself where such socio-political issues are surfaced to viewers - reality TV. This genre is often dubbed 'poverty porn' [127], as it often focuses on people who are supported by state welfare. The recent proliferation of this genre has led to concerns that prime-time media experiences are exacerbating misconceptions, and stifling critical debate, around major societal issues such as welfare reform and poverty. This genre of TV capitalises on othering and stigmatisation of people in order to increase viewing figures. This is well documented in previous literature, which demonstrates that audiences of reality TV are often not critical in their viewing [217], therefore this is a problematic genre, viewed by uncritical viewers, which works to exacerbate and entrench existing stereotypes and myths about those claiming state welfare [221].

Therefore this domain is concerned with the consumption of socio-politicised TV media, and the redesign and reconfiguration of this process, to structure and encourage reflection by the viewer/user. To this end, I conducted two studies, focused on different aspects of the TV consumption and discussion process. The first study presents a smartphone app, Spotting Guide, that allowed users to codify, in the form of short tags, patterns or behaviours they notice within a programme. This enacted critical viewing by the users through their use of the app. It was deployed as a

set of three workshops, which also involved participants using the app in their own homes. The second study iterates on Spotting Guide. I produced a smartphone app, Screenr, that enables collective co-selection of programmes, and the subsequent critical co-viewing of the selected programmes. It was deployed in-situ with 12 participants for the period of one month, who used the app on their own device in their own homes, with little involvement from any researchers.

5.1 Existing Practices

In the previous chapter I focused on social media. Here I will shift the focus towards the *process* of second-screening, encompassing the interaction with data streams and apps on the second screen, as well as what is being shown on the TV. Second-screening is an interesting context for a digital prototype as it offers the ability to connect multiple spheres of viewers [53], which extends the connection of those co-viewing the TV in the same room. Cesar et al. [53] characterise the interactions made possible through second screening in four ways: users might control the broadcast using the second screen; they might transfer content between devices and the television; broadcasts might be enriched with user-generated content; and users might be encouraged to share personalised content with other viewers.

The possibility of enriching TV content with second-screen applications has been recognised by the broadcast industry. For instance, HBO's "companion apps" are designed to be used in conjunction with series such as Game of Thrones [118]. These apps allow users to see extra context-relevant information about the show as they are watching it, including the backstories of characters and recaps of previous episodes. Similarly, Nandakumar & Murray [176] present experiences in evaluating a companion app for a long-form US television series. This application was designed in the form of an annotated 'story-map', which provided the backstory for characters within the program to scaffold viewers' first viewings of the programme. Basapur et al. [13] incorporate the idea of social sharing as a means of providing content for second screening. Their application, FanFeeds, was designed to allow users to provide time-coded commentary (e.g. text, video, URLs, etc.) alongside TV programmes to be shared with like-minded friends. As the authors note, this form

of user-generated content lends a feeling of a live social event amongst TV users, they also note that users drew on friends' comments as talking points, to engage in 'offline' conversations or conversations on other platforms beyond the broadcast of the programme. Moreover, users also reported that the sharing aspects of the platform encouraged deeper consideration about their commenting practices; users expressed desires to produce content that their friends would find interesting, funny or otherwise engaging, and doing so required reflection on the imagined audience reception for such comments.

While second screens have been shown to heighten certain forms of engagement with TV programming, they do also raise specific challenges. In both Anstead et al. [8] and Basapur et al. [13] it is noted that designers have to mitigate the potential difficulties users may find in managing their viewing across multiple screens without disrupting or sacrificing the "liveness" of the event itself (e.g. pausing the broadcast to post or read comments). Second-screening applications present an opportunity for enriching broadcast content through the encouragement and support of users in sharing information in specific, supported, ways.

Specific Interactions for Critical Second-Screening

The previous studies demonstrate how people already engage with second screening and socio-political discussion around reality TV, and the potential for bespoke applications to support more critical viewing and co-viewing of broadcast media. However, these studies also demonstrate the difficulties users may experience in navigating around and unpicking user-generated second screen content, due to the ways in which these broadcast media interrelate socio-political issues and TV entertainment. The question is therefore: how can interaction design facilitate and encourage critical reflection whilst second-screening? I have identified two key mechanisms from the literature that present opportunities to do so.

Leaning Forward / Leaning Back Vaccari et al. emphasise the capability for existing second screening practices to blur the distinction between "lean-forward" and "lean-back" practices. They note that second screening blurs and complicates

the relationship between information seeking and *“relatively passive, information-reception practices classically associated with broadcast media"* [237, p.1044]. They explore the effects of using interaction design to place a greater emphasis on more active “lean-forward" practices noting that amongst their study participants the “relatively active" practices of live commenting *“contributed to a statistically significant increase in individuals; propensity to engage in a range of political activities"* [237, p.1055]. Hence, they conclude that while second screening has the potential to influence political or civic engagement, it may only be able to do so if it supports non-passive involvement with the broadcast media rather than just serving as an *“additional source of information"* about it [237, p.1055].

Social Tagging Previous research has explored the idea of social tagging as a non-passive mode of engagement with information. Ames & Naaman [7] explore tagging practices around two photo-sharing applications, ZoneTag and Flickr, noting that users simultaneously used tags to convey opinions around the content of media and to improve image searchability for themselves and others. They identify two functions of tagging: organization and communication as well as two modes of sociality: tagging for-the-self and tagging for-others. Users’ tagging practices were noted to fluidly move between all types, with images often containing tags that fitted all four criteria. Overall, tagging was shown to facilitate critical thinking about how tags would be received and understood by others. Kammerer et al. [130] and Nelson et al. [179] studied two applications - MrTaggy and SparTag.us - which use social tagging to append contextual metadata to web pages to be used in enhanced web searches (MrTaggy) and more easily select relevant information from within web pages (SparTag.us). Both studies demonstrate that the activity of information-seeking with tags encourages users to engage differently, and more critically, with information corpora. In Kammerer et al. [130], users of MrTaggy reported being more engaged in exploring search results with social tags than participants using standard search engines. Similarly, in Nelson et al. [179] SparTag.us users visited fewer URL sources, relying more on tags to describe and make sense of provided information. This was especially so when those tags had been produced by friends. Hence, across both applications, the tags themselves act not merely as navigational tools but are a resource for engaging more deeply and critically with content.

Information seeking with social tags therefore seems to have the capacity to support users in sense-making work and to interpret and situate their own responses alongside the responses of others. However, Bodoff & Vaknin [28] also show that the act of tagging information in itself is important in terms of encouraging users to think critically about information. Their study investigated how semantic priming only occurred when users had themselves tagged similar information previously [91]. In other words, users reflect on information, or its tagging, more after they have viewed other people's tags or have produced their own tags to assist others in sense making.

There is, therefore, a design opportunity for new modes of interaction and spaces for discussion around reality TV, where moral and critical reactions to reality TV can be promoted. The aim would be to encourage critical viewing that unpacked and explored underlying production decisions, the framing of a socio-political issue, or consideration of the viewer's own views and beliefs.

5.2 Spotting Guide

Motivated by this previous research context, I set about designing and deploying a smartphone app, Spotting Guide, that would facilitate viewers of reality TV in critiquing the programme. To do this, Spotting Guide encourages users to identify, categorise and tag patterns and tropes that they see within a TV programme. It is also premised on encouraging less passive and more active second screening activity, utilising interactions such as 'spotting' and 'tagging', as opposed to merely passively watching the programme. This, according to the research discussed previously, has the potential to support lasting engagement with the socio-political issues emerging from broadcast media. In the following sections, I describe the design and implementation of Spotting Guide, the study design, the findings, and a discussion of the findings.

Designing Spotting Guide

The design and initial implementation of Spotting Guide was conducted together with Tom Schofield, Senior Research Assistant at Newcastle University. We worked

collaboratively to reach a design that would support critical viewing of reality TV. This resulted in an initial prototype of Spotting Guide. Tom built a concept interface for Spotting Guide. After this, I took over the technical implementation and all future design decisions.

The Spotting Guide app was inspired by traditional paper-based “spotting guides", such as the Michelin I-Spy spotting guides popular in the UK [163]. Aimed at children, these books motivated readers to look for specific objects, such as train signals or bird species, and record them in their spotting guide, culminating in a reward from Michelin for completing the guide. The books were well circulated, and the “spotting" genre covers many subjects. As such, the concept of “spotting" occupies a place in the cultural imagination of the UK and is familiar to many people, as noted by [19]. Motivated by this, it was envisaged that a digital spotting guide could be created for reality TV. Users would be required to look for and record interesting patterns they see in a programme, hitting a ‘+’ button next to the desired category (Figure 5.1). If users identify something interesting that is not in their spotting guide at all, they can type a new category in, and use this to spot future occurrences.

The design of Spotting Guide is intended to allow users to identify and spot patterns of reality TV production that portray people in a negative or stereotypical way. At the technical level, the spotting experience begins with the user synchronising their activity with the beginning of the broadcast via a ‘start’ button, allowing the app to be used with live and recorded programmes and those viewed through a digital “catchup" service. The app was developed as a web app, using HTML5 and JavaScript. The app was optimised for use on smartphones and tablets with testing across a variety of common web browsers. Spotting data was stored centrally in a secure manner, which was subsequently printed to aid discussion in the workshops.

5.2.1 Study Design

Participants were recruited via poster and email advertisement distributed around Northumbria University and Newcastle University, in Newcastle upon Tyne, UK, as well as local cafes and bars. To maintain discussion group sizes and to aide scheduling, two groups were created. A total of 12 participants took part in the study, divided into two groups of five and seven participants. Participants ranged between 18 and

Figure 5.1: Mock-up of Spotting Guide being used as a second-screen

35 years of age, and comprised six females and six males in total, and per group: two female three male and four female and three male. All were familiar with reality TV, and used social media. Participants were compensated for their time with a £40 voucher. The evaluation was conducted over a three-week period, with one workshop per week. Each workshop involved a discussion, along with short activities in relation to the Spotting Guide. Between each of these workshops participants were asked to complete a “homework" exercise. These involved using Spotting Guide whilst watching television. Participant experiences and the outcomes of their ‘spotting’ activities would then be discussed in the following workshop. Overall this formed three workshops, and two homework activities per group.

- **Workshop 1:** Comprised a short discussion to contextualise the research, as well as explore the participants’ knowledge and feelings towards reality television, and specifically othering and stigmatisation. Participants were also introduced to the app, and through a series of training exercises, shown how to use it.
- **Homework 1:** Participants watched an episode of Benefits Street, at home, whilst using the app at the same time. Benefits Street was selected as it is a well-known UK reality TV programme, which follows the lives of people living on state welfare in a deprived area of the UK.
- **Workshop 2:** Began by inviting participants to discuss their experiences of watching and using the app, followed by a group exploration of their anonymised ‘spotting’ data.
- **Homework 2:** Participants were asked to select their own programme to use the app whilst watching. They were provided a “blank" app, with no pre-created categories, and asked to create these themselves.
- **Workshop 3:** Comprised a short two minute presentation by each participant of the programme they selected, and their experience. This then flowed into a group discussion about their homework task supported by participants’ data, and a final reflection on the overall experience.

I was assisted in the workshops by Kiel Long, Senior Research Assistant at Northumbria University. During the workshops, all of the discussions were audio recorded and then subsequently transcribed. Inductive thematic analysis was conducted [39] on this data. Two other researchers (Gavin Wood and Scarlett Rowland) assisted as second and third coders. I produced an initial codebook based on an initial pass of the data, I then discussed this codebook with the other two researchers, until we all agreed on the codes. I then used this agreed codebook to recode the data. At this point, we collaboratively clustered the codes to create a thematic structure for the data.

5.2.2 Findings

Below I present the themes from the analysis, which are grouped around two overall themes of Acts of Doing and Acts of Thinking.

Acts of Doing

This theme comprises participants' views of how they used Spotting Guide and how it affected their viewing experience. This is divided into three sub-themes: Making Spots, Accuracy of Spots, and Spotting While Paying Attention.

Making Spots Discussing the initial "seed" spotting categories created by the research team for Homework 1, participants had mixed views of their value: *"'Dogs' - I never felt they were particularly important to the narrative, it was just there"* [P5, G1]. When creating spotting categories, using things that were visually prominent in the programme was common: *"...a really messy bedroom with just thousands of cosmetic products littering the floor, and picking up on those sorts of visuals"* [P7, G2] and *"I made* [a category] *for swearing, which was a very bad idea because I had to press it... a lot"* [P4, G1]. The interaction to spot an instance of a category was described as easy, compared to the live-tweeting process: *"Something will prompt you to tweet, and then you focus on that... it will take longer than tapping a button for an existing* [spotting category]*."* [P2, G1]. The interaction was also described as game-like: *"The closest parallel I can think of, is US or British elections where you've*

got buzzword bingo... something you tick off and first person to get a line wins" [P2, G1].

Accuracy of Spots Throughout the study participants expressed a desire to be consistent when they were spotting and creating categories: *"The problem of wanting to redo it. By the end of 25 minutes I created a category... if I went back and watched them again I think more things would have been spotted"* [P4, G1]. Difficulty deciding whether to make a spotting category was another concern of the participants: *"I kind of regretted that I hadn't made a spotting category... I realised something had repeated and I hadn't made one... you wanted to go back and respot it"* [P3, G1]. Over time, participants observed that the meaning behind a spotting category they defined changed over time: *"I had one saying drug use that came from cigarettes, then I think there was a point that* [I] *started classifying cigarettes as drug use"* [P8, G2]; and: *"A spot that I created during the intro because I thought it was going to come up a lot, it was 'war analogies' and I ended up repurposing it for things like violence"* [P6, G2].

Spotting While Paying Attention Several participants described the high cognitive load required to watch a television programme, critically analyse it and integrate this analysis into an app: *"It definitely came in fits and starts* [...] *where suddenly I was going swearing, crime, rubbish - all within about ten seconds. ... I was trying to do too much at once..."* [P7, G2]. Specifically, it was noted that remembering all of the spotting categories created was difficult: *"You've got a long list in your head, it's hard to keep track. Hang on - do I need to create another spot or is this already covered?"* [P2, G1]. Participants noted how attention was often split between the app and the television, notably due to the amount of work required to use the app: *"You are concentrating on typing... I must have had 20 spotting categories altogether, so I was constantly doing something whilst absorbing information"* [P1, G1]. Participants also noted how their attention would be split between the audio and the visual elements of the programme: *"I'm much more of an auditory person, so I was focussing on what was being said"* [P4, G1]; and: *"...focussing more on what was being said by the characters rather than seeing it because it was very difficult to see*

and spot at the same time" [P7, G2]. Finally, issues with attending to spots and the shows at the same time were further influenced by the pacing and editing of the programmes being spotted: *"It's very slick - there's music overlaid, there is commentary, there is what people are saying, then there is whatever visual happens to be on screen. Then switch to something going on in the background whilst that person is still speaking"* [P2, G1]; and *"That was a major thing at how fast paced these programmes are and you don't realise it. Then there's someone with a cigarette talking about money, then you want to make another* [spotting category]*"* [P3, G1].

Acts of Thinking

This theme encapsulates participants' wider reflections around issues presented in reality TV, as well as how the process of using the app provoked or inhibited reflection. This is divided into three sub-themes: *Reflecting on Spots*, *Interpreting Content* and *Sharing of Spots*.

Reflecting on Spots During the workshops, participants mentioned they would not normally watch reality TV type programmes such as Benefits Street, but two had seen an episode of it before. They reflected on how Spotting Guide could encourage them to think about the programme production: *"One way of challenging conceptions about othering is to get people to pay attention to what it is the programme makers are doing"* [P1, G1]. These sentiments were tied into an overall view that the app could promote mindful viewing of TV - being aware of the content in the show, as well as how and why it has been selected, edited and shown in a particular way: *"There are two types of watching the app has brought out for me. The turn your head off watching and then there is being analytical"* [P3, G1]. One participant humorously remarked: *"It's going to ruin my viewing pleasure from now on... it won't be mindless anymore. I'm just going to deconstruct everything"* [P10, G2], while another noted: *"talking about and analysing all our data [...] allows you to go 'There - I've been manipulated!'"* [P7, G2]. Participants discussed the nature of ad breaks, and how this could be used to reflect on the content in Spotting Guide: *"*[During ad] *breaks you think 'I'll make a cup of tea', but it would be a different kind of dynamic if during*

the ad break you were reflecting on the data. It kind of gives you time to do that" [P5, G1].

Interpreting Content Participants suggested that the pre-existing prejudices and beliefs of a user might be reflected in how they use the app, through the kinds of things they would spot: *"If you gave the app to someone who was predisposed negatively to* [claiming state welfare]*, they would be going 'oh my god, how terrible!' They wouldn't be thinking through behind the scenes"* [P2, G1]. The subjectivity of how a spot is interpreted, or noticed in the first place was discussed: *"If you have one person from Vote Leave and Vote Remain* [Brexit] *watching exactly the same thing, they would come up with different responses... we all have certain viewpoints."* [P5, G1]. Participants openly reflected on how their preconceptions affected their programme choice and the data they created: *"I deliberately chose a show that I knew would be playing to particular stereotypes with the people in it. So I started trying to spot what I thought were particular instances of the stereotype"* [P7, G2]. Further to this, reflecting broadly on second screening activities to provoke thought about reality TV, participants noted if the ideological views of the second screening activity are too "alien" to their own views, they may disengage from the app: *"For example spot every time the Conservative government do something lovely - if that's too alien then they would just not engage"* [P1, G1].

Sharing of Spots A sense of curiosity was shown by participants, towards what the rest of the group were doing: *"I would be interested to look at their spots, as we were watching the same programme, just to see how they map onto each other, the similarities and the differences."* [P1, G1]. They also described how showing their homework spots in the next workshop affected their spotting practice: *"In a way I knew I was going to be judged... the* [other participants] *are going to look at my spots"* [P3, G1] and *"I like sharing my opinions, but I would be more careful if I knew I would share spots, so would try to make it as neutral as possible"* [P8, G2]. However, a few participants expressed concern that sharing spots could negatively affect other users by conforming to the majority opinion: *"Exposing people to the comments of others could potentially force them to reflect on their own comment if it is different... my concern is the danger of the majority"* [P1, G1].

5.2.3 Discussion

I will now present a short discussion of the results from the Spotting Guide deployment, with a more comprehensive discussion presented in Chapter 7.

Split attention between screens

Participants noted they had to split their attention between the screens, with some describing the spotting going in *“fits and starts"* as they rapidly switched their attention. It was also clear that participants had to use their memory to keep track of the spotting categories and their meaning. In the Spotting Guide, the context of the activity is within the programme itself, and therefore, whilst it does not involve an investigative process to understand extra context, it does demand more attention to the programme and effort on behalf of the user to translate that into the app. The capacity to support critical reflection around co-viewing of broadcast media may, counter-intuitively, also rely on provisioning for less critically-oriented activities. As we saw, the participants used the ad breaks as a moment of rest, or a time to reflect on what they had just seen, and they noted explicitly the ad breaks provided what they saw as much needed rest from the activity. Put differently, it may be infeasible to ask users to perpetually “lean forward" into non-passive engagements when second-screening, and some periods of “lean back" relaxation are seemingly required in order to potentially deepen and maximise engagement with the application in a broader sense [237].

Designing for friction

There is an important tension between viewing the TV programme and interacting with the app. Spotting Guide participants expressed that using the app was an intense interaction. The design of Spotting Guide demonstrates the effect different interactions have on the second screening experience with a view to guiding the type of reflection and discussion that occurs towards the programme.

In the context of civic participation, Korn and Voida [135], discuss the ways in which “friction", conscious design decisions that make interactions less seamless, which take a non-neutral stance or position, can be incorporated into the design of

digital systems to encourage deeper critical engagement and responses from people in contexts and situations where passivity is the norm. Whereas Basapur et al. [13] and Anstead et al. [8] express concerns about the fluidity and smoothness of engagement between viewing across a TV broadcast and an application, I view the intensity, and friction of the Spotting Guide interaction as positive. The spotting process itself provoked viewers to engage more critically than they might normally in examining the production techniques and portrayal of characters, as it highlighted *“two types of watching [...] turn your head off watching and then there is being analytical"* [P3, G1]. Participants were forced to engage more deeply with the content of the show by making trade-offs between creating and adding “spots" against watching the programme; and developing strategies where they focused more on sounds and speech rather than what was shown, as noted by P7 *“focussing more on what was being said... it was very difficult to see and spot at the same time"* [P7, G2]. While undoubtedly hard work, this aligns with Korn and Voida’s principle that *“designs for friction want to cause trouble. They do not want to help you; rather they place little obstacles in your way"* [135, p.8]. Therefore, with the difficulty and intensity in spotting portrayals in reality TV, Spotting Guide carved out *“space for reflection in the residue between activities"* [135, p.8].

Allow users to re-define what tags mean to them

It is clear that participants were reviewing their own “work" and resultant data during the study. This also extended to their own definition for the behaviours, patterns and sentiments they were classifying, which aligns with Bodoff and Vaknin’s [28] depiction of tagging as a way of non-passively engaging with information, and as a means of encouraging users to reflect on their opinions around the information to be tagged. For example, one participant reflected how their spotting categories *cigarettes* and *drug use* began to overlap during the course of the programme, which then prompted a deeper reflection of their own views towards cigarettes and drug use. This is a clear example of a user reflecting on the content of the show, prompted by the use of Spotting Guide. Future designs may include a user-editable “description" field accompanying each category or more thoroughly through a diary-type interaction where a user can record how their meanings have changed over the

course of use. Design considerations may also include the capacity for users to remove, redact or rename tags, or to provide multiple tags for content to reflect the conflicting opinions that may be experienced by individual taggers.

The audience as a resource for critical thinking

The nature of the Spotting Guide study encouraged its users to think about the audience reception of their tags and spots, which evidences Kammerer et al.'s [130] and Nelson et al.'s [179] claims that the social tagging process is not simply about annotation. Rather, tagging encouraged more critique and reflection around the users' own standpoints on an issue and how the views of others *"map onto each other, the similarities and the differences"* [P1, G1], forcing the users to think about what their tag would mean shared publicly. Participants also openly admitted they self-censored their input into the system because they *"knew I was going to be judged... the* [other participants] *are going to look at my spots"* [P3, G1]. This is unsurprising, given that existing work suggests that an increasingly critical engagement with broadcast media is premised on a level of interaction with them that goes beyond the passive reception of TV shows and their related user-generated content. As Baumer [15] notes, when confronted with a morass of different and competing standpoints on a socio-political issue (such as with unfiltered social media feeds), it may be difficult to even identify and situate your own opinions within the wider debate. For designs that utilise social sharing of data, this is significant because it shows, in the context of polarising reality TV, that social tagging moves beyond simple annotation of content towards critical/reflective thinking around the user's own viewpoints. Hence, balancing users' capacity to self-censure against the core mechanic of openly sharing genuine opinions with a real or imagined public is a key consideration for second screening applications designed to support critical/reflective thinking.

5.3 Screenr

Spotting Guide demonstrated how a different mode of interaction whilst second-screening can support viewers in critical reflection on reality TV. Spotting Guide used 'social tagging' [7] of on-screen behaviours as a means to actively engage

users in the critical viewing process. Participants knew they would be talking about their tags in front of the focus group the following week, and thus incorporating a social element to 'tagging' encouraged taggers to consider the potential audiences for those tags. The Spotting Guide study also showed us that there is a balance to be found between attention-demanding 'lean-forward' activities such as tagging, and more passive 'lean-back' activities such reading and reflecting on your own content.

Motivated by these findings, in this section I report on the design and deployment of a mobile application, *Screenr* (Figure 5.2) built to explore co-selection and critical co-viewing of reality TV. Screenr allows a group of viewers to collectively vote to decide upon a live television programme to watch together each week; the app then supports the critical co-viewing of that programme during its live broadcast. Whereas Spotting Guide was exploratory in nature, and therefore used a workshop based methodology, I wanted to understand how a critical second-screening application would function in the real world, as a part of participants normal TV viewing practices. I deployed Screenr with 13 participants in situ over a period of 4 weeks. During each week participants were asked to collectively decide upon a show to watch, to simultaneously view the show in their own homes, and to collaboratively comment upon, tag and discuss the show's content. The app was intentionally designed to support critical co-viewing and constructive reflection on TV content; I therefore deliberately chose reality TV content for the study, which is known to generate predominantly uncritical reactions in viewers [41].

First I document the design process for Screenr, highlighting what I learned from the Spotting Guide study and how this influenced the design of Screenr. I report on the findings of this study and explore ways in which the co-selection of programmes can be facilitated over a period of time, and how live, critical co-viewing can be supported and encouraged by apps such as Screenr. The study contributes the following insights: the mechanisms required to coordinate co-selection of programmes for critical viewing, managing the attentional demand of critical second-screening, and the design decisions to encourage different types of critical reflection.

Figure 5.2: Mock up of Screenr usage in the home

5.3.1 Designing for Critical Co-Viewing

Based on the motivation to introduce sociality to the critical viewing process, via co-viewing, I set out to design a system to support the critical co-viewing of 'everyday' reality TV programmes. The previous study Spotting Guide focused on encouraging critical reflection in a second-screening application by individual viewers, but did not explore the social aspect presented by co-viewing with others, facilitated through a second-screening app. As can be seen in the Spotting Guide results and discussion, participants were already keen to know what other people were saying (and indeed, this is part of the motivation for online backchannels, as discussed in Chapter 4), as well as being mindful to temper their input into the app, knowing that others would eventually see what they had spotted. Also, in the Spotting Guide participants already discussed how they went through an iterative process of re-defining what tags meant to them as the programme was going on, and I surmise here that this tagging process, which evidently involved them reflecting on the meaning of the

words and phrases that formed the tags, would work as a social tagging service, prompting reflection on the tags, and the assumptions and framing being made by the other users.

To this end, I designed and implemented a mobile second-screening app, *Screenr*, to be used for critical co-viewing of live TV programmes. Through the design, I hoped Screenr would i) promote 'closer-readings' of TV through asking viewers to summarise on-screen talk, action, and sequences through tagging, spotting and importantly, chat; and ii) to make these readings visible to other users to foster dialogue around show content. To encourage a variety of views and differences between users, the tagging process was left open to interpretation to reflect this. We can also see that users of Spotting Guide had varying levels of cognitive load, some finding it difficult and others producing many tags. Making this process social possibly benefits the whole group of viewers - those who find the process easy will become content producers, whereas others who prefer to focus on the screen can engage with the content being produced. Based on these opportunities presented by the Spotting Guide study, I was motivated to design a system that could encourage critical co-reflection as part of a co-viewing experience. The following section details the specific design decisions taken to structure co-selection of programmes, and to facilitate and encourage critical discussion during live broadcast.

5.3.2 Designing Screenr

The Screenr interface is shown in Figures 5.3, 5.4 and 5.5. The app consists of four main features: i. the co-selection of TV shows for weekly co-viewing; ii. a live interface for tagging on-screen patterns and behaviours; iii. a chat interface for open, free-form discussion around the programmes; and iv. a scrapbook for private reflection.

i. Co-selection and voting: It is imperative that a mutually agreed programme be established in order for co-viewing to be successful. Yet there are considerable socio-technical challenges when designing for co-viewing of live broadcasts. People have unique, often unpredictable schedules and may need to coordinate with other household members who also have commitments. As such, coordinating multiple users to co-view at the same time around programmes that may lie outside of their

usual preferences becomes a significant scheduling problem. I selected live broadcast television for Screenr as, despite increased platform choice in recent years, industry reports (see [16]) indicate that live viewing is still the most popular way to view TV in the UK, ensuring the app would fit into a normative viewing environment. TV schedules are typically released 10 days before broadcast which means coordinating multiple people for co-viewing quickly and decisively. It would have been relatively simple to dictate to users a pre-determined schedule; however I wanted participants to have some control to select shows, and encourage group discussion about the topics and programme choices available each week. In Screenr I adopted an open public voting system to coordinate co-viewing. Therefore, I leveraged co-selection mechanisms, such as public voting, which have been proven to be effective tools for gauging and disseminating voting preferences of a group [244].

In the app, a digital TV guide displays details of the available programmes which users can browse through, and cast a single vote. Each programme is displayed with an image, a short summary, when and which channel it will be broadcast on, and who has voted for the programme so far (see Figure 5.3b). Users press a button below the programme they wish to vote for. Upon voting, their vote is reflected alongside the programmes listed in the TV guide, e.g. “User A has voted for this". Voting is open between Sunday until Wednesday morning; these days were chosen so as not to overlap with the predominant reality TV broadcast schedule (Wednesday - Friday). On Wednesday voting is automatically closed and the programme with the most votes is selected for viewing that week, with the programme selected randomly when there is a tie in votes. All users are emailed with the outcome of the vote and the selected programme’s details. A log of each user’s vote is publicly displayed on the home screen (Figure 5.3a), along with a tally of votes for each programme. Both of these mechanisms intended to expose, amongst others, the scheduling preferences of the group as a whole, and provide a degree of steering towards a mutual programme to co-view.

ii. Live interface: This is used when watching live broadcasts, and functions as a textual tagging interface for critical viewing, building on Spotting Guide, where tagging provides a sense-making process for the live programme. There are three key functionalities available to users: *Tagging*, *Spotting*, and *Importing* (Figure 5.4a and b). *Tagging* allows users to create short tags (up to 40 characters) that encapsulate

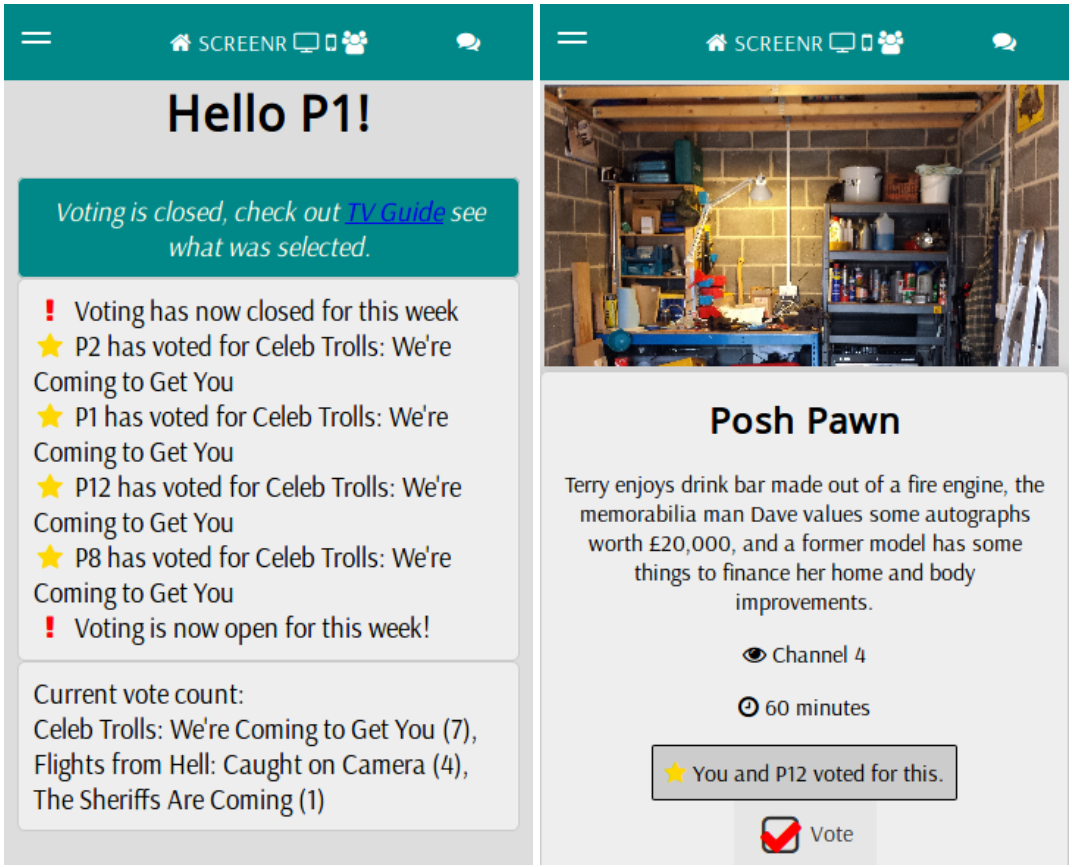

Figure 5.3: Screenr interfaces: a) Home screen showing voting and scheduling information b) TV Guide for voting on programmes

an on-screen behaviour or pattern in the programme. Tags created by the user are displayed in a grid with a counter beside. *Spotting* involves tapping on a tag to mark that it has occurred in the programme and at that point in time. Spotting can be performed on any tag within the user's live interface. The amount of times a tag has been spotted is displayed by a counter alongside the tag. *Importing* allows users to copy another user's tag and bring it into their own live interface. Figure 5.4b shows the real-time list of all other users' tags that are available for import. Imported tags can then be used in the same way as the user's own tags, and are denoted by a purple background, compared to a green background for the user's own tags. If a user creates a tag that already exists, the existing tag is imported into the user's live interface. Tags can be deleted from the live interface at any time, and re-added

should a user wish.

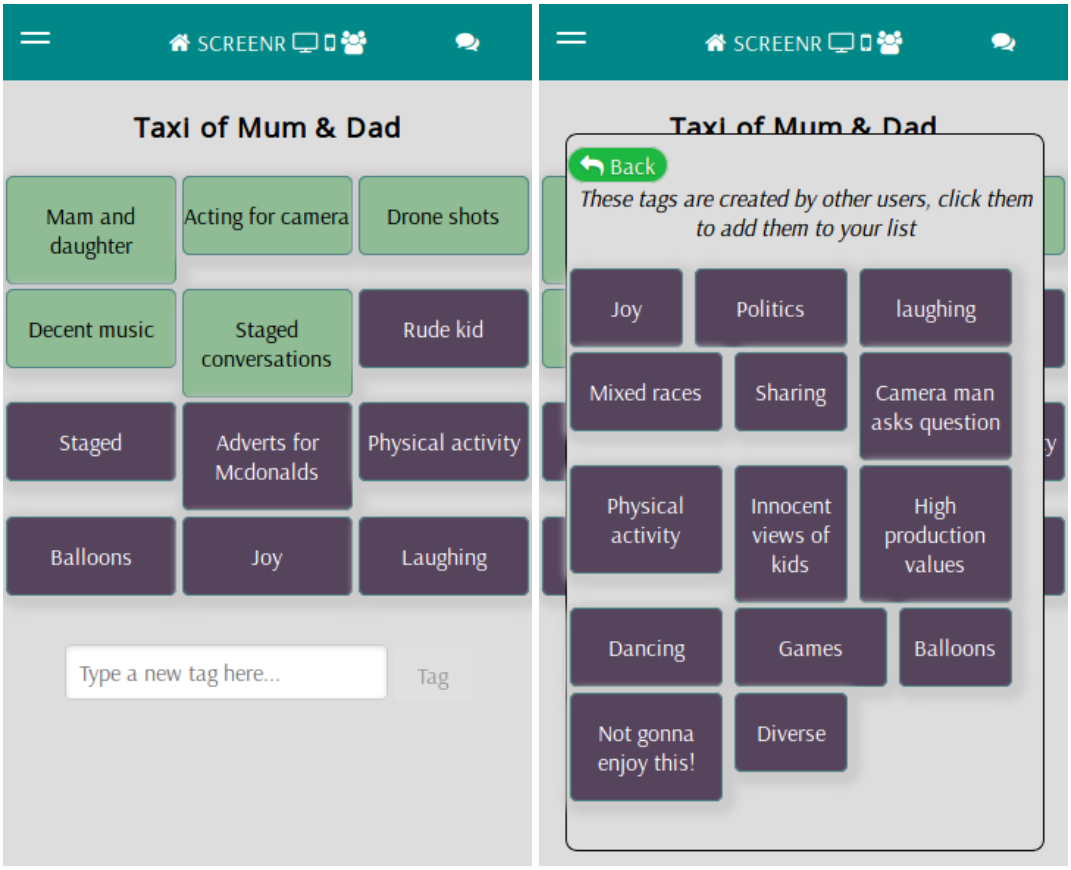

Figure 5.4: Screenr interfaces: a) Live interface showing tags that can be spotted b) Importing tags from other users.

iii. Chat interface: In order to capture discussion in a variety of ways, I designed both a public discussion space, and private reflection space within the app. All public discussion in the app takes place in a single location, the chat interface (see Figure 5.5). This is an instant-messaging style interface that allows all users to share messages with each other as part of a single, continual conversation thread.

iv. Scrapbook: While the focus is on collaborative viewing and discussion, the scrapbook allows the users to share their private thoughts with the research team, capturing any comments a user may not wish to share with the group.

Throughout the application, rather than displaying user's names, pseudonyms were used. As per my prior work with Spotting Guide, and as demonstrated by

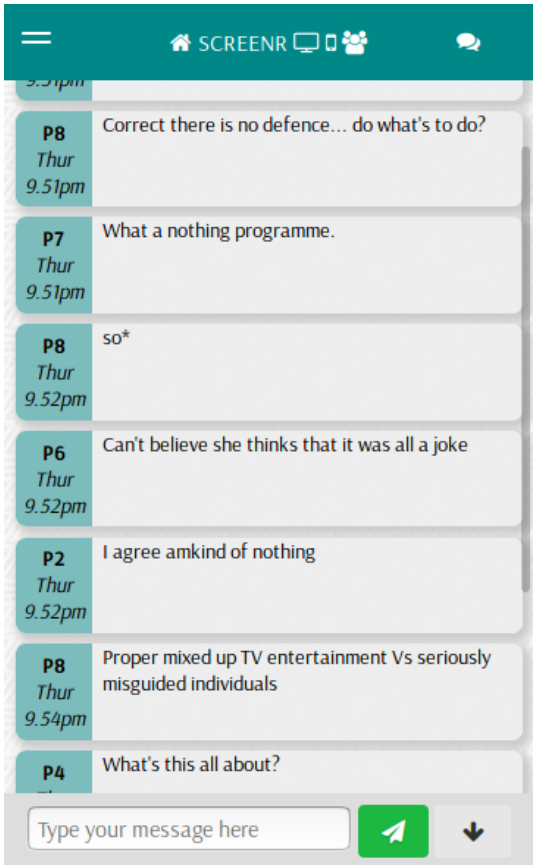

Figure 5.5: Screenr's instant-messaging style chat interface.

Gorkovenko et al. [107], I expected this would allow users to contribute to discussion around potentially sensitive and politically charged topics without revealing their identities. At the same time, the pseudonym provided them with a unique moniker with which to be identified by other participants in subsequent conversations. Pseudonyms were selected at random from a list of animal names. The *Home* screen features as the central hub for the app (see Figure 5.3a) which shows recent system activity, such as user votes, vote opening, and vote closing. This is accompanied by a display of the current vote counts. The Home screen displays contextual reminders to users who have not voted, and it displays the details of the voted programme following the vote closing. Automated email reminders are also sent to notify users when new programmes are released, when they have not voted, and when voting has closed.

Screenr Weekly Cycle

Screenr operates on a two-part weekly cycle running from Sunday to Saturday, detailed below.

Part 1: Voting to watch a programme On Sunday, new programmes become visible within the app. The voting system also becomes open simultaneously, automatically closing on Wednesday morning. Once visible, users are required to cast a single, non-transferrable vote. They do this by browsing the TV guide (Figure 5.3b), and hitting the vote button beside their desired option. At this point it is the intention that users will reflect on the programmes available in terms of which they think would make a good critical co-viewing experience for the group. After voting, users are prompted to discuss their voting choice in the chat interface with the rest of the group or in their own personal scrapbook. The winning programme is decided when voting closes, with users notified by email the outcome of the vote, reminding them what and when they will be live viewing.

Part 2: Live viewing of the selected programme This part starts at the date and time of the selected programme. Users are directed to use the *Live Viewing* interface (Figure 5.4), where they are able to create, import, and spot tags. Participants are requested to switch their TV to the correct channel. When other users create tags, a small message animates on the live interface (“New tags!") to notify the user. Users can press on this message to view the latest tags, and import any to their live interface should they chose. Throughout the programme, users are able to import, create, and spot tags in their interface. Users are able to switch to the Chat interface at any time should they wish to participate in discussion.

Technical Implementation

Screenr is implemented as a web app, primarily to avoid cross-platform compatibility issues and to make it available on a wide variety of mobile devices. The front-end of the application was created using the AngularJS framework, which incorporates HTML, CSS and JavaScript. The back-end of the system was built with Node.js, using MongoDB for database services. An API was created with Node.js and Restify, which allowed the front-end to retrieve and update information via HTTP requests.

5.3.3 Study Design

To deploy Screenr, and in line with the recommendations of previous work [8, 107], I conducted an in situ deployment to explore how co-selection and critical co-viewing would work in the home environment. Notably, deploying in participants' homes would allow participants to become familiar with the app, as well as vote on and watch programmes from actual TV schedules. The study was designed to fit a normative TV viewing schedule in the UK context, using widely available channels at prime time. Participants used their own devices in their own homes with little interference from myself.

At the commencement of the study, participants were asked to individually attend a session to orient them to the study, as well as ensure the Screenr app worked correctly on their device. During the study, participants were asked to use Screenr each week to vote for a programme and then subsequently watch the selected programme at the allotted time. This was done each week for four weeks. In the third week of the study, participants were contacted by telephone for a 15 minute interview to ensuring they were having no technical problems with the app and to check that they were engaging with the process. In the final week of the study, three focus groups were conducted each using between 3-5 participants per group, lasting approximately one hour each. Each focus group was divided into three parts. The first part concerned the overall experiences of the participants, such as their experiences of co-viewing and which parts of the app each participant preferred to use. The second part utilised “decks of cards" made up of all the programmes available to vote on throughout the study, through which I led participants through each week of voting and asked them to reflect on their voting choice, as well as discuss the programmes as a group. The final stage utilised printed tag clouds from each co-viewing session. Participants were asked to talk about their own co-viewing experience, and the tags they created and saw throughout the programme. I was assisted by Gavin Wood, Senior Research Assistant, and Scarlett Rowland, Research Intern, from my research lab, in the interviews, telephone interviews and focus groups.

Participants

I worked with a recruitment company to recruit participants, and following the early withdrawal of one participant, the study involved 13 participants. The inclusion criteria was for participants to be over the age of 18, owning a recent smartphone, and regularly watching TV between 6pm and 11pm during the week. All participants lived in the north-east region of the UK to ensure they could travel into Newcastle upon Tyne for the initial interview and final focus group. Participant ages ranged from 25 to 55, with eight identifying as female and seven as male. The majority of participants were experienced in using common technology such as smartphones, laptops, and desktop computers. All participants owned a recent smartphone (less than 3 years old), were familiar with its use, were daily viewers of TV, and regularly watched reality TV programmes. I did not seek participants who specifically viewed reality TV, nor did I seek participants with specific political viewpoints. Over the duration of the four week study, the study involved participants giving up 12 hours of their time. Participants were compensated for this, with a total of £150, at an hourly rate of £12.50, based on suggested best practice [86]. To my knowledge none of the participants knew one another. At the start of the study, participants completed a 15 minute interview in order to understand their experiences of second-screening, reality TV, and othering and stigmatization on TV. They varied on a number of dimensions and these are captured below:

Reality TV viewing practices All participants had seen reality TV. Some participants (P4, P5, P6, P7, P8, P11, P13) were enthusiastic viewers of reality TV, others (P1, P3, P9, P10, P12) were occasional viewers: *"I went off them for a little while, then I seemed to be watching more and more again"* [P10]; *"Big Brother, Love Island, I don't watch any of that, but I'll watch all the police things"* [P3]. One participant was a reluctant viewer and watched simply because another family member had chosen it: *"Not necessarily through choice, it's cause what the wife's got on the TV"* [P2].

Critical Viewing The majority of participants were relatively uncritical viewers of reality TV and would be drawn to the programmes because they provided entertainment that was easy to watch. However, a couple were critical of the genre in

general: *“I don’t see much constructive in it [...] poking the hornet’s nest to make something happen"* [P10], or claimed a more thoughtful approach to viewing: *“I am quite analytical when I’m watching these programmes. They may be trying to portray someone as particularly violent and superficially you might firstly see the aggression in that behaviour, but I’m trying to unpick anything else they’ve done."* [P8].

Second-Screening All participants, except one, had previously engaged in second-screening unrelated to the primary screen (e.g. browsing social media or conduct unrelated Internet searches). Furthermore, the majority of participants had used their second-screens in relation to the primary screen and these activities included fact-checking and using social media to view and/or join in a conversation or debate about a particular show. One participant used group messaging between friends as a backchannel for a specific programme: *“Between friends I use WhatsApp [...] That just seems to have got everyone involved."* [P1].

Choosing Programmes for Screenr

The initial criteria for selection was that programmes should be broadcast in the evening, to broadly match the participant’s schedules, and that programmes should be reality TV genre. In order to reduce votes becoming too thinly spread and to provide steering for the co-viewing process, a maximum of four programmes were offered to choose between each week. The reality TV genre offered a wide range of programme topics available for critical co-viewing each week. All programmes were selected from the UK government run Freeview service [90], which includes the most commonly viewed TV channels [26].

Data Collection and Analysis

An array of usage data was collected within the Screenr system. All chat messages were recorded, including the author username and time and date stamp. Similarly, entries to the scrapbook and programme suggestions were recorded. Within the tagging interface, the following information was recorded: creation of new tags, *importing* someone else’s tag to your tagging interface and *spotting* a tag (own or imported). Votes and voting choice each week were also collected for each user. In

total this yielded 378 chat messages, 897 tags, 1105 imports of tags, 775 spotting instances, and 50 votes.

Each entrance interview, telephone interview, and focus group was audio recorded, and subsequently transcribed. An inductive thematic analysis method was used to analyse these transcripts, as described by [39]. I initially identified codes within the data, with Gavin Wood acting as a separate coder, separately coding a subset of the data. We both discussed our codes and, when we reached agreement, I re-coded the data with the new codebook. Gavin and I then collaboratively clustered the coded data into a set of themes.

5.3.4 Findings

The findings are divided into three sections. First, to give a sense of the system use, a detailed example of usage during one of the live viewing sessions is provided. This is followed by the overall patterns of use across the duration of the study. Finally, I discuss the outcomes of the thematic analysis, describing insights into the ways that Screenr provoked critical reflection as well as some of the social processes at play. I use an anonymised notation throughout for participants (e.g. P1), with focus group sessions being noted as FG1 through 3, and weeks abbreviated to Wk1 through 4.

Screenr Usage Example

In order to demonstrate how the system works overall, the following is a vignette of the second-screening experience, taken from week three of the study. *The Taxi of Mum and Dad* was a reality TV programme aired in the UK on Channel 4 in August 2017. Billed as *“eavesdropping on conversations between parents and teens"*. The programme received 7 votes and thus was chosen to be watched by the majority of the group. The programme depicted conversations of 8 families in cars from fixed cameras pointing inwards towards the car’s occupants. The programme was fast paced, showing a conversation from one family for approximately 10 minutes, with an overarching story linking the conversations throughout the programme. Two minutes before broadcast, one participant created the tag *“Sitting waiting"* [P5]. Once the programme got underway, some participants used the chat interface to reflect on

their choice of programme: *“I’ve got this programme wrong. I thought was going to be about mam & dad taxing [sic] their kids everywhere"* [P9, Chat].

In one scene, a number of on-screen characters are seen eating fast food in the car. Within seconds of this, two users had created tags: *“Everyone likes junkfood"* and *“Advert for McDonalds"*. This second tag was then imported by two other participants into their own tagging interface. This sentiment was continued in the chat: *“Have a conversation round the dining table man... not in the car having a macci dees"* [P8, Chat]. A short montage followed, in which many of the programme’s characters posed to take selfies, creating a flurry of tags within seconds of each other: *“Selfies"* [P9], and *“Selfies, so hip"* [P10]. Tags were also used to identify production features, such as the tag *“Rude son/rebellious music"* [P10] created in reference to one of the characters whose argument was accompanied by rock music.

Throughout their experience, the participants questioned the production values of the programme: *“Why on earth is this set in a car?"* [P8, Chat], which was responded with: *“No idea P8 thinking the same"* [P9, Chat]. One popular tag was created by P7: *“Acting for camera"* which was imported shortly after creation by P12 and P2, with all three users spotting this tag throughout the programme. Another participant shared their feelings at intervals during the co-viewing: *“Not real life totally staged, not what I had thought it would be!"* [P6, Chat] and *“A lot of it must be put on for the cameras"* [P6, Chat]. In this example, tags described on-screen objects and patterns, with chat providing a longer form of co-discussion. The following section provides a more in-depth analysis of this kind of participant data, taken across the whole study period.

Overall System Usage

Four rounds of voting were conducted over the four week study, with no ties. Wk1-3 participants had four programmes to vote between, with Wk4 containing three. Votes were generally spread across all programmes, with every programme receiving at least one vote each week (see Table 5.1). The programmes selected tended to be on Wednesday or Thursday, starting between 7pm and 9pm. In total, 13 participants used the system for a period of 4 weeks. Of those, one engaged very little

with the system throughout the study only casting votes, and whilst the remaining participants used all features of the app.

Overall, usage tended to focus on Sundays, when participants were notified of new programmes released onto the app, and on Wednesday or Thursday when the chosen programme was typically aired. Outside these times, the chat interface was used to raise technical questions (*"What is the scrapbook??"* [P2, Wk1]), or reflect on the voting process itself (*"I voted for the winner at last"* [P3, Wk3]). An overview of participant activity over the four weeks can be seen in Table 5.2. We can see that participant engagement remained constant, if not slightly increased, over the course of the study, with almost all users voting. The programme viewed in week 2 was 30 minutes long (instead of 60), which thus resulted in roughly half as many interactions on Screenr. Week 3 is the most prominent in terms of engagement, with the largest amount of social interactions (e.g. chat messages, importing tags). It is evident that participants differed in the extent to which they had previously engaged in critical co-viewing and this played out in their overall use of Screenr. Figure 5.6 shows a representation of all 13 participants who contributed to the study, showing their contribution to each aspect of the reflective process. For example, P2 predominantly imported tags and created tags, thus their engagement with co-viewing was focused on tag curation and creation, as they imported tags heavily from the rest of the group, as well as creating tags for their own and others' use. P8 was focused very heavily on chat discussion, with a small degree of importing of other users tags, meaning they used Screenr mainly for social discussion, with tags being sourced from the rest of the group rather than creating their own. P1 in the centre has contributed evenly to all aspects of Screenr. The slight alignment to the left-hand side of the graph indicates they engaged in slightly more tagging and on-screen spotting of tags than anything else.

Through these three example participants, we can see that P2 contributed heavily to the tagging process by creating and using others' tags, with P8 contributing the most messages to the chat discussion and solely using others' tags. P1 is more balanced in their contributions, and slightly more focused on the TV screen and spotting task. Participants expressed these desires to use specific parts of the app: *"I did read the conversations through the ad breaks, then at the end of the show I'd put*

Table 5.1: Programmes voted on each week.

Week	Programme	Votes
1	Old People's Home for 4 Year Olds	5
	The Secret Life of the Holiday Resort	3
	Nightmare Tenants: Get Out of My House	2
	GPs: Behind Closed Doors	1
2	Fake Britain	7
	Nightmare Tenants, Slum Landlords	2
	Body Fixers	2
	Traffic Cops	2
3	Taxi of Mum and Dad	7
	No More Boys and Girls	3
	Nightmare Tenants, Slum Landlords	2
	Don't Deport Me, I'm British	1
4	Celeb Trolls: We're Coming to Get You	7
	Flights from Hell: Caught on Camera	4
	The Sheriffs Are Coming	1

a comment on how I felt the show was ... The chat for me was reading what other people put." [P11, FG3].

Results of Thematic Analysis

Analysis of the qualitative data - tags, chat logs, notes from participants and transcripts of interviews and focus groups - captures the kinds of critical reflection that resulted from Screenr use. From the data, six themes were constructed.

Table 5.2: Screenr usage statistics over four weeks

	Wk1	Wk2	Wk3	Wk4
Total Votes	11	13	13	12
Tags Created	297	110	299	191
Tags Imported	384	110	443	168
Tags Spotted	318	59	182	216
Chat Messages	78	42	136	122

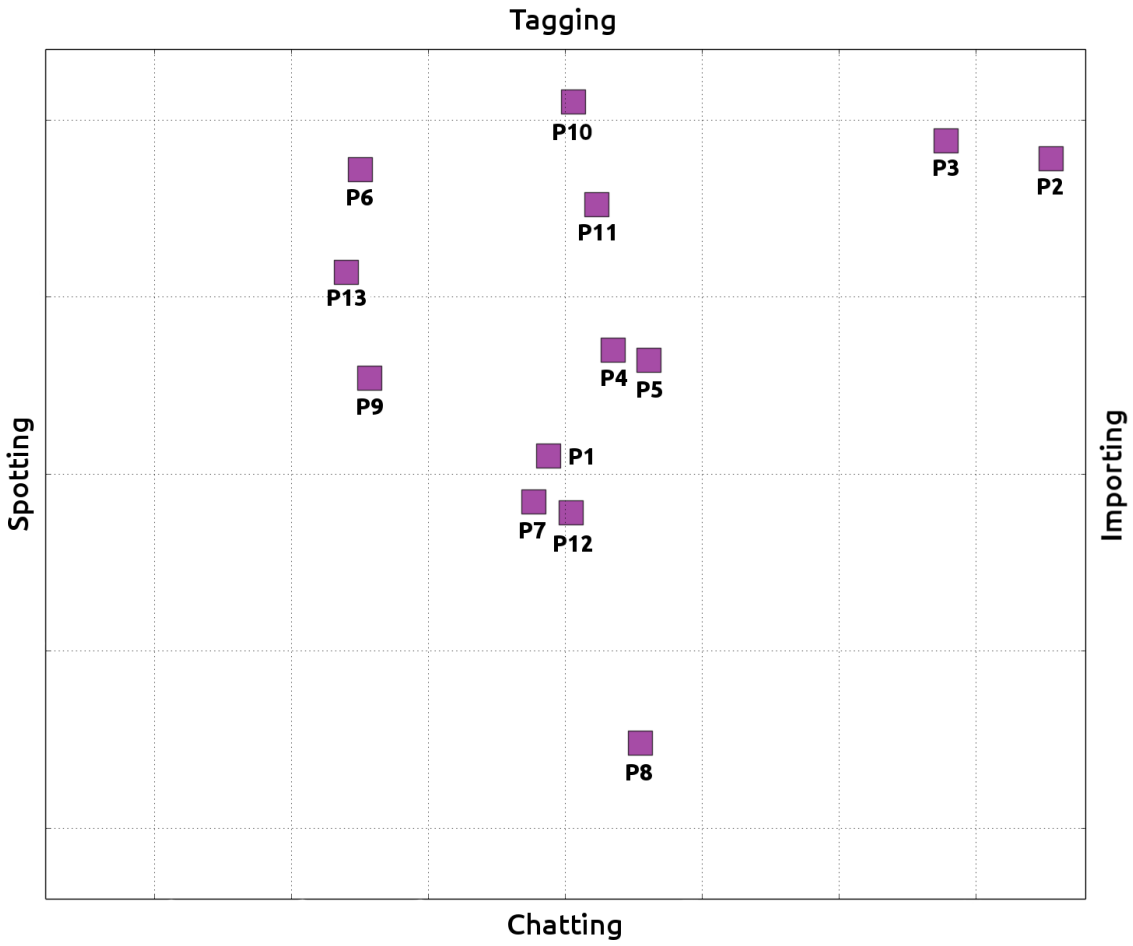

Figure 5.6: Balance of contribution throughout study for each of the 13 participants.

Critique through tagging Through analysis of tagging behaviours, we see that participants were often critical viewers of programme content and questioned either the “claims" made by the producers or the portrayal of people in the shows. For example, the programme in Wk1 depicted a social experiment with elders and infants. In the segment that focused on statistics and the social experiment, a large number of tags were created to interrogate these claims: *“False hope?? Revert back?"* [P2, Tags], *“How long would it last?"* [P4, Tags] and *“No negative feedback"* [P10, Tags]. Critical viewing was often associated with the depiction of controversial social issues, where participants would use chat to discuss what they considered may be provocative. In week 3, two participants had a conversation about this: *“Here we*

go with the class divide banter" [P8, Chat, Wk3], shortly followed by: *"Is it [the programme] othering children or their parents, or the parent child interaction?"* [Chat, P8, Wk3]. This was followed by *"Think the othering is of parents looking clueless"* [Chat, P7, Wk3]. Participants also challenged the production choices, tagging obvious visual features: *"Close up of feet?"* [P6, Tags, Wk1], and *"Culturally diverse"* [P10, Tags, Wk3]. Critique was also directed towards portrayal of people. The tag *"Assume it's illegal immig"* [P10, Tags, Wk2] was created, and then imported twice by other participants, and within 5 minutes, one participant replied directly to the tag creator: *"I agree with whoever tagged that we automatically assume it is illegal immigrants"* [P12, Chat, Wk2].

Social affirmation and influence There was often a strong social bonding component, which I describe on the following pages, whereby participants affirmed each other's reflections and discussion points: *"Agreed P8, can't ignore when they're about your kids!"* [P12, Chat, Wk4]. This often took place when participants used the chat function to discuss social issues. For example, in Wk1 discussion covered old age, and state/societal responsibility: *"Scandalous... maybe it's best if some old people aren't stimulated, it costs too much"* [P8, Chat, Wk1], and *"It does make you think though if it would have the same effect if it was a council home with kids from low income/benefits attending"* [P13, Chat, Wk1]. These appeared alongside the tags: *"Policy change afoot"* [P8, Tags, Wk1], *"Society should do more"* [P2, Tags, Wk1] and *"Do OAPs need CRBs?"* [P10, Tags, Wk1].

In the focus group discussions, participants recognised two sources of influence. First, aligned with comments made during the chat sessions, they acknowledged that they were sometimes manipulated by the programme producers: *"But in [the programme] the way they portrayed the northerners... the bias, the music... it categorised them"* [P7, FG2]. Others recognized that they were unable to know to what extent they were subject to manipulation: *"Is that othering on the part of the television programme, or is that just actually the reality of a lot of dodgy characters, I don't know."* [P10, FG1].

Second, the participants recognized that Screenr itself was a source of social influence, noting that they used the group as a sounding board to assess the extent to which their own views were reflected in the tagging activities of others. Some

likened their use of Screenr to other social media systems such as Facebook: *"It's like a 'like' on Facebook, using another* [person's] *tag"* [P7, FG2], *"I tended to look at the other tags... 'Ooo that was a good way of putting it, why didn't I put that.' My likes of other people's tags would have been in the ad breaks."* [P4, FG3]. People also recognized that the tags and chats of others could influence their own judgement: *"It was... people highlighted and I was thinking, 'Am I thinking this or am I letting people's thoughts influence me'."* [P9, FG1].

Reactionary forms of engagement Not all tagging was as considered, however. Sometimes participants clearly displayed unthinking reactions to content, sometimes taking the form of anger directed at the people portrayed in the programme. An example is found in the chat conversation in week 3 and 4: *"Jumped up opinion of self"* [P2, Tags, Wk3], *"36 and a granny, why am I not surprised"* [P4, Chat, Wk3] and *"Vile people"* [P9, Tags, Wk4]. At other times, strong feelings of outrage would grow, partly as a function of tagging and chat as part of the second-screening process. During the fourth programme, which focused on people extensively engaged in online abuse (a.k.a. trolling), one participant discusses their own views: *"The 'extreme trolls' are so damaged... why would you do that!"* [P8, Chat, Wk4], which they follow up shortly after: *"I have totally 'othered' the trolls" ... "I'm getting so angry.." ... "I said before I wouldn't want an apology... I'd want a scalp!"* [P8, Chat, Wk4]. Intertwined, but not directly responding, another voices their thoughts towards their views: *"Surely these Trolls are mentally ill? 40 messages a day??"* [P7, Chat, Wk4], which is later followed by *"I'm completely othering the Trolls. Never heard of this happening, its horrendous."* [P7, Chat, Wk4]. There is a certain irony here, given that the system is developed to expose and reflect on such processes, to the fact that participants are stigmatising the people depicted in the programme and a form of social escalation takes place in the process. The implication here is that there may be behaviours such as flaming (personal attacks) or Othering that emerge during second-screening and we should anticipate this kind of unwelcome response.

As is typical of social media exchanges, participants could become quite heated about certain issues during the tagging and chatting process and various forms of social escalation and 'flaming' would take place. Some participants recognized this:

“it’s a faceless platform, there’s no worry about anybody seeing who’s saying it" [P10, FG1]; *“I think that’s dangerous, and I think that’s [...] why I would never fully engage with an app like Screenr [...] I would much rather have the courage and the convictions in a room full of people than the courage and convictions on a faceless app"* [P8, FG2]. As a consequence, some would carefully moderate their own input: *“I was a bit politer than I normally would be, I say what I think, but in a more guarded fashion in this group."* [P7, FG2]. Others made a purposeful decision to withhold their own views or regulate their input within the app: *“I don’t want to have those sort of issues when I’m eating my dinner and relaxing [...] so I wanted to be a bit more passive, so I just ignored that one"* [P2, FG2], and *“I didn’t know whether I wanted to chat with everybody on that topic [immigration], if I’m honest... I thought it could end up like... heated. I thought ‘yeah let’s not go there’, so that’s another reason why I didn’t pick that one"* [P6, FG1].

Social accountability and viewing context Participants also felt the need to be socially accountable. For example, they sometimes watched and voted on the talent programmes and in such cases would use Screenr to account for their choices, particularly if they were adrift of group norms. Sometimes they would explain the social context of the vote in terms of own their family dynamics: *“The only reason I voted for it is I’ve got two teenagers and I thought this is going to be real life, that’s gonna be great."* [P3, FG2]. At other times they would actively involve other members of the family in the on-screen discussion: *“I was asking my husband and I was commenting on the comments that were on there and saying, ‘They’re all saying this’, or ‘Somebody’s saying that, what do you think of that?’ so I was involving him in the conversation. But I mean I don’t think... he might’ve influenced what I was commenting on but only if I agreed with it and made me think about something else"* [P8, FG2].

Limitations of tag length The word limit for tags, implemented to keep tags short and pithy, caused difficulty when trying to encapsulate on-screen behaviour or thoughts on sensitive issues: *“it was finding the tags were quite short to express really what I was trying to say a lot of the time, so a little bit more room for the tags would’ve probably been a bit useful."* [P10, FG1] and *“It was quite hard, trying to think of different*

words to sum it up, because you had a certain amount of letters that you could put in [...] how can I shorten that into a better word?" [P4, FG3]. The character limit to tags did not present difficulty for all participants, with one method being to use tags and chat for different types of discussion: *"Tags for me, it's just summarizing what's happened in the programme* [...] *the* [chat] *was your feelings towards what was happening, what was being said"* [P4, FG3] and *"With the chat you can obviously say more, so it might give more context to what the tags are aiming at. Between the two you'd probably get a good picture."* [P10, FG1]

Attention and critical co-viewing Splitting of attention was experienced by all participants throughout the study, with attention being directed to one of the screens based on a number of factors. One participant said that using Screenr changed their TV viewing experience from a passive viewing experience into an active second-screening experience: *"I couldn't just switch off and watch the programmes because I felt like I had to be tagging things and contributing to the discussion"* [P12, FG3]. However, the level of this active engagement was also individually moderated by the participant's interest in the programme topic: *"I felt happier when I was commenting on something I was interested in* [...] *watching something that I wasn't interest in, it was more of a chore."* [P8, FG2] and *"If they're heavy subjects people especially on a night time, week on week, can't be bothered to sit down and work that bit hard in watching something and trying to work out what they think about it."* [P10 FG1].

The quantity of data being generated during live viewing within Screenr could lead participants to focus on the app at the expense of the programme: *"I felt I missed quite a lot of the programme whilst I was reading what everybody else was saying and then making my own comment, and then I was like 'oh I've just missed like that whole 30 second segment'"* [P4 FG3]. Similarly, others found they became more engaged with the discussion over time, as it changed and grew: *"At the beginning of each programme it was okay,* [..] *but once it started racking up, and you're trying to read what people are saying, then you're trying to think, put your bit in then somebody puts something else in it changes your"* [P3 FG2], with the attentional demand leading participants to feel unable to engage with others in the room: *"The first week my husband tried to say something and I was like 'Shh, shh!'"* [P6 FG1]. Given the attentional demand from the second-screening process, participants leveraged

advert breaks, for those programmes that had them, to keep track of discussion: *“you wanted to make sure you were catching everybody else’s tags.* [...] *somebody had said in the chat bit about not liking adverts but thank god for adverts."* [P9, FG1], *“I felt like I caught up on the conversation when the adverts were on"* [P13, FG3], and *“I used [the ad breaks] to have a rest. I just put my phone down for a couple of minutes.* [...] *Nobody really spoke during the ad breaks on the chat as well so, a bit of calm."* [P12 FG3].

5.3.5 Discussion

Screenr was designed as a means to encourage co-selection and critical co-viewing of reality TV. The findings reflect how I hoped participants would use Screenr, as they demonstrate users conducted close readings of the programmes viewed and engaged in a spectrum of critical discussion, whilst consuming and interacting with each others’ data. I now discuss some of the pragmatic design lessons learned from the in situ co-viewing experiences before going on to consider the lessons learned about how to foster critical co-reflection.

Designing for Real World Co-Viewing

The findings highlight a number of pragmatic considerations when designing for co-viewing in users homes, which I unpack in the following sections.

Provide diverse avenues for user engagement The results indicate a need for flexible forms of engagement as users differed in their preferences. Some predominantly used the tagging functionality - effectively *packaging* and labelling small pieces of the programme - and would use the chat sparingly and thoughtfully. Others relied on the chat interface, contributing quick comments but also sourcing tags from the group, through the importing mechanism. These show profoundly different patterns of use, yet both support co-viewing and critical reflection.

These different usage profiles also map to social roles. A user may move between a creator producing tags for the group, to a consumer importing tags, then flipping to a socialite focusing on the chat interface. This echoes existing work around social roles in tagging, such as Thom-Santelli et al. [232] who describe

how some users may publish and produce content, whereas others may explore existing tags. Therefore I argue the need for flexibility in designing co-viewing systems to cater to a wide range of social roles and encourage diverse participation.

Open voting for co-selection When facilitating a co-selection process, use of open and public voting mechanisms can be beneficial in a number of ways. Co-viewing of live broadcasts has numerous tensions, most notably the requirement for all users to be present at the same time, and the relatively short forecasting of TV schedules. As such, it is imperative that systems designed to facilitate the co-selection of programmes for live co-viewing leverage techniques to encourage and guide voting. At a base level, a hung vote or indecisive vote is not conducive to a self-sustaining co-selection process, and would thus require an external arbitrator, who is possibly unknown to the co-selection group. When designing for co-selection of programmes, voting as part of a group helps to build investment in the co-viewing task.

Fostering Critical Co-Reflection

Baumer [14] proposes that designs to encourage critical reflection should facilitate the reflective process of *Inquiry*, through creation of designated inquiry spaces, and fostering group discussion. Screenr responds to these design strategies by allowing co-viewing users to contribute to a discrete, live, social tagging interface, and a central overarching group discussion. The group discussion through the chat feature fostered a wide range of critical reflection, such as on the production choices and values of programmes, the wider issues raised by the programmes, but also reflectively on participants own values and judgements.

Building group criticality As can be seen in the results, participants are discussing Othering and production techniques on their own terms, which they are not equating as negative but have still isolated this act within their conversations. This mode of discussion supported criticality among participants; this criticality was co-created together, rather than being explicitly directed by an external expert critique. As such, by reconfiguring the boundaries between ‘expert’ and ‘TV viewer’, Screenr

allowed them to discuss, in their own terms, their critical reflection process. Critical viewing is often oriented towards programmes already identified as problematic (cf. [41]), however the results show that critical tools can be turned to a variety of everyday programmes (within a specific genre), sometimes highlighting manipulations in production (Wk3), or deeper societal issues (Wk1). This wider discussion presented by the group, somewhat a 'critical collective', addresses some of the underlying issues with reality TV as a genre.

Engaging and disengaging with criticality It is clear in the context of reality TV that there is a tension in relation to programme topics. In the study, when presented with voting options for programmes that centred around topical UK political issues (e.g. immigration, gender identity) some participants avoided these programmes, attributing this to not wanting to engage in any kind of discussion on these issues with others whilst watching. This raised broader questions about the applicability and frequency of use of second-screening apps for criticality. Previous work has noted how viewers may well be aware of online discussion around a programme, but purposefully chose not to engage in it [106]. I speculate that whilst 'lean-forward' engagement with second-screening can result in useful critical engagement with the programme being viewed [237], there are some programmes where viewers would rather 'lean-back' and not engage in any discussion.

Liveness as opportunity for group criticality The Spotting Guide study demonstrated that users find the process of tagging during second-screening attention demanding, and within these designs a degree of passive engagement with the programme is required. The social tagging features of Screenr speak to this design consideration. As the results show, participants were able to browse tags created by others to either rephrase or reframe a topic they were interested in, or endorse and adopt a topic created by another participant by importing it to their interface. At a system level this indicates criticality can be widely sourced from the other co-viewers, which in turn reduces the attentional load for users. Furthermore, participants used the advert breaks when they were available as a means to relax from the co-viewing activity. This mirrors how participants used the advert breaks during the Spotting Guide study, where participants noted that advert breaks provide a

design opportunity for deeper reflection. The social aspects of Screenr - the importing of tags, and the chat discussion - leverage this design opportunity and provide users with a space for reflection, passive or otherwise, in between and at the end of programmes.

5.4 Summary

Overall, this domain of inquiry tells us about the design and deployment of two digital prototypes, that introduce criticality to viewers of reality TV, leveraging a second-screening approach to do this. At the outset of this chapter I detailed existing practices of those who second-screen, as well as the research being done with the design and development of novel second-screening applications. Furthermore I isolated a number of techniques that could be used to foster criticality among viewers. I presented the first study, Spotting Guide, which structured viewers in critiquing the production values of the programmes by recording 'spots', in the form of tags. The study, operationalised through a series of workshops, demonstrated that criticality can be fostered with viewers in this way, but that challenges are present, most notably the requirement for viewers to disengage from the critical process, and the desire for a social element. Motivated to address these opportunities, I iterated on this to develop a second digital prototype, Screenr, which introduced social elements to the original design of Spotting Guide. Furthermore, wanting to understand how such a system could operate in a real world setting, I designed Screenr with programme co-selection in mind, so that viewers themselves would agree on the programme to critique. I deployed Screenr in situ for them to use for one month. The results show Screenr's varied second-screening activities allow users to fit the system around their preferred mode of critical reflection, and as such form an ecosystem of content producers, adopters, and observers. The results also allow me to highlight a number of challenges and opportunities for the designers of critical co-viewing applications, such as the importance of providing distinct spaces for inquiry and using group discussion as a means for group sensemaking. A full exploration of the results of both of these studies can be seen in the discussion (Chapter 7). The following chapter focuses on the final domain of inquiry of this thesis, Conversational Interfaces.

Chapter 6

Domain Three - Conversational Interfaces

The focus of this chapter is on conversational interfaces. As I noted in the introduction, these are a novel technology that is in the early stages of mass market adoption, and they present opportunities for users to engage with socio-political topics in powerful ways. Indeed, smart home devices such as Amazon's Alexa are already bringing ethical and moral questions to the fore, as they are reconfiguring the way data and algorithms are engaged with within the home, as I will explain. In contrast to the previous chapters, in this chapter I focus on the opportunities that such a novel technology presents for the design of agonistic interfaces. I start by discussing the current technologies and conventions being used to design and develop conversational interfaces, and focus particularly on VUIs and smart home assistants. Taking up the opportunities I highlight, I document the design and deployment of a smart home assistant, Spkr, that broadcasts diverse viewpoints on current socio-political topics into the home environment.

6.1 Existing Practices

Advances in natural language processing and digital communication systems have, in recent years, introduced conversational interfaces and conversational agents into a variety of contexts. Whilst the idea of the conversational agent is not new, having been popularised by Weizenbaum in the 1960s [248], however it is only with the

aforementioned technical advancements that have led to their integration into systems and platforms used by large numbers of users. Klopfenstein et al. provide a detailed history of conversational interfaces throughout the years, citing the different nomenclature that has been used, such as chatterbots, virtual personal assistants, and chatbots [134]. They note that conversational interfaces have historically undergone personalisation, where they become conversational *agents*, with a distinct name and some form of embodiment. There is also an emerging discussion as to what defines a *conversational interface*, and whether they are truly conversational in the same way that we refer to human conversation, or rather they involve *talking* as part of the interface. See Reeves [200] and [201] for a more in-depth discussion of this. Hence, I use the term voice user interface (VUI) when describing Spkr and related devices, and chatbot or bot when describing text based interfaces, as I feel this more accurately captures the conversational abilities of the devices, in line with Reeves' critique.

Social media and instant messaging platforms have introduced their own environments that allow developers to create their own conversational agents. These are often known as bots or chatbots, a contraction of software robots. Bots perform a variety of functions, and have been touted bots as *"the new apps"* [204], with many companies developing and deploying bots to improve their customer experience [181]. However, their ability to operate continuously in an autonomous manner presents challenges in some contexts. The detection of problematic "spam bots" on social networks is an active area of research [245], and more recently the way bots may influence political opinion and decision making is being explored [120]. Exploring the ethics of bot designs, Maréchal [153] examined the ways the terms and conditions of different platforms guide the creation of bots, specifically noting how different platforms facilitate or limit aspects of human rights.

Virtual personal assistants are a further extension of the conversational interface concept, which has been facilitated by technological developments, such as access to machine learning and voice recognition services online [134]. Their popularity has been facilitated by the smartphone, which has become somewhat ubiquitous. These virtual personal assistants were embedded in smartphones, starting with Apple's Siri, and being followed by Google's Assistant and Microsoft's Cortana [134].

Based on the expanded capabilities of voice recognition technology, these assistants are able to perform a wide range of tasks on the smartphone, using natural language. With the boom in Internet of Things (IoT) technology in the home [46], virtual personal assistants have found their way into household objects, most notably the smart speaker, also known as the smart home assistant. The Amazon Echo is a popular smart home assistant, which embeds the virtual personal assistant Alexa[1] as a core part of its interface, with the user relying on a VUI to operate the speaker. The Amazon Echo and other smart home assistants perform functions such as playing music, providing a weather forecast and managing to do lists [218]. They use voice recognition to detect when someone in the house is addressing the device, such as “Alexa" or “Hey Google", which uses microphones to listen to the household discussion. There are growing concerns over the presence of such devices where *“users trade privacy for convenience"* [141]. More generally, virtual personal assistants are also being explored as a way to encourage users to trust a system, for example with self driving cars [140].

Being “Pushy"

Push notifications are a common part of the smartphone environment, originally created by Apple as a means to send messages from a server to the user interface [160]. Push notifications are now a feature on desktop applications, as well as through websites. They are commonly used to send messages to the user, notifying them that, for example, a new email has arrived, or an Instagram follower has liked their photo. Thinking more broadly, push notifications effectively reach out to users without any request - rather than the user checking their email themselves, the push notification tells them, *there and then*, that they have received a message. Here there are analogues with human conversation, as after all conversation is a two way communication, with either party able to start or stop the conversation at any point. Current chatbots and VUI design conventions mean that they only “speak when spoken to”, and do not push themselves on the user without being invoked first.

[1] https://developer.amazon.com/en-GB/alexa

The convention to “speak only when spoken to" is dictated to a degree by the platforms on which the chatbot or VUI are built. Bots that utilise social media platforms do so through an API (Application Program Interface), which provides access to various functions such as searching, liking, friending, posting, and so on. The usage of the platform is restricted through one or more sets of terms and conditions (T&C) created by the platforms to regulate usage in order to fit their overall strategic goal, such as to provide relevant content for users, and to reduce offensive content. Using Twitter as an example, there are two sets of T&C governing the deployment of bots, one for general usage of the Twitter platform, and a more specific set of “Automation rules" [2]. This latter set of T&C gives some specific guidance to developers, such as not to post "spam" content, and not to directly contact users unless requested to do so.

Text-based chatbots that are built specifically for smartphones are able to use whatever systems are accessible, such as the push notification system, where they leverage this to their advantage. Woebot[3], a chatbot for cognitive behavioural therapy (CBT), is an example of a chatbot app that uses push notifications to remind users to interact with it. This is driven by the theoretical underpinnings of CBT, which is usually formed of a set of sessions over a number of weeks. Woebot uses push notifications, framed as if they are from Woebot itself (“Hi Tom, it’s been a while since we last spoke...") to encourage the user to maintain the CBT process. While the user can simply ignore the notifications, using the notification acts as a nudge [227] towards a positive health outcome - the continuation of their CBT.

The current implementations of smart home assistants such as Amazon Echo and Google Home do not allow the device to be activated without first being triggered by a *wake word*, such as “Alexa" or “Hey Google". As a developer, the API for Alexa does not allow you to initiate a conversation without a speech *trigger*. At a technical level, the speech recognition system within the Echo device triggers the software routines in the Alexa assistant when a person speaks to the device[4], and thus it is not possible to activate the Alexa without it being spoken to first. Developers

[2]Twitter: Automation Rules. Available from: https://support.twitter.com/articles/76915

[3]https://www.woebot.io

[4]https://developer.amazon.com/en-US/docs/alexa/ask-overviews/understanding-how-users-interact-with-skills.html

have been keen to explore this, but have also reached the same conclusion[5]. Some apps that have been validated by Amazon are allowed to perform conventional push notifications, and therefore speak without being spoken to, for example to notify that a smart doorbell is ringing [82]. However this is for opt-in services that the owner has purchased/activated themselves. Some users have experiences Alexa speaking unprompted, as in March 2018 some users experienced their Alexa laughing randomly, with no human prompt [85], and some users report similar occurrences with Google Home[6]. These are acknowledged as errors with the speech recognition software.

Given the existing privacy concerns around smart home assistants, developments are already pointing towards devices that do not listen all the time, but only when spoken to [162]. No doubt, this is an important step for privacy, but I argue there still remains opportunity to build smart home devices that *push* their content onto the user, rather than only responding when spoken to. Allowing agents, whether using a VUI or a text-based interface, to reach out to users, within reason, provides an important, enriching and useful experience. It is clear that a balance must be found, given how disruptive, and annoying, such a feature could become. As a counter point, enabling corporations and governments, to harry, remind and cajole us on a large scale through social networks and smart home devices, may lead us towards a *"Kafka-esque nightmare"*, where such systems exert huge socio-bureaucratic influence over our lives - something analogously described by [71] when discussing the automated anti-fraud telephone systems historically used by banks.

There are examples from the research context that use VUIs and virtual personal assistants in a pushy way, for a variety of different purposes. Gaver et al.'s Energy Babble is a radio-like device that reads out information related to energy conservation. Taking it's design cues from the radio, it plays content without being prompted, and allows users to respond to it with their own questions [101]. The Social Printers, proposed by Gorkovenko et al. [107], do not use speech, but do use a pushy characteristic. They were smart home devices that contained a small thermal printer (as used to print receipts for goods) that would print out prompts and questions for

[5]https://iot.stackexchange.com/a/883
[6]https://support.google.com/googlenest/thread/1482775?hl=en

discussion by those in the living room, sent by others participants or the research team. The print outs were timed to coincide with televised political debates, that all those in the household would be watching. The act of producing a printed question, undoubtedly got the attention of those in the room (e.g. *"what do you imagine may happen in the case of leaving the EU?"* [107, p.2271]), and would thus initiate action, such as getting up to tear off the receipt and look at the question.

Taking the idea of pushy smart home devices to the extreme, the (theoretical) *Nag-baztag* proposed by Kirman et al. [133] is an example of radical behaviour change using negative reinforcement, which was enacted by making the device pushy. The Nag-baztag device would monitor the energy usage of the household and provide personalised feedback for members of the household, as well as taking actions based on this. This would not be prompted directly, but would be triggered by actions in the household, such as putting the kettle on. In order to get household members to reduce their energy consumption, the Nag-baztag *needed* to be pushy, so as to stop ambivalence on the part of the user. Taking actions that affected their environment was a means to get their attention, and punish them for their wasteful behaviour.

In the following study I leverage the opportunities presented by smart home devices utilising pushy behaviour by building a digital prototype, Spkr, that introduces those around it to diverse socio-political topics that are being contentiously debated.

6.2 Spkr

A range of factors are responsible for reducing the diversity of viewpoints encountered in the online public sphere. The presence of highly partisan media outlets, social factors such as selective exposure to news topics and social homophily (see [2, 12, 20, 98, 114]), as well as the design of the systems themselves [191] means that citizens are increasingly likely to encounter political viewpoints that confirm their own. Whilst this online polarisation is less pronounced than in offline media [103], Socio-political discussion online is nuanced, and is influenced by a variety of sources, including non-news outlets [172, 173]. Around certain sensitive topics, an overwhelming negative opinion can effectively drown out other, more diverse, viewpoints [41].

As discussed in Chapter 2, diversity of viewpoints is considered part of a healthy democracy [80, 175], and thus the narrowing of exposure to diverse opinions in the online public is problematic. Indeed, work is ongoing in this area to broaden users' exposure. Munson et al. experimented with inserting oppositional viewpoints directly into the user's news feed [172], whilst in a later study Munson et al. evaluated a browser plugin that would demonstrate to users their own news bias based on their reading history. Others have explored building devices to encourage reflection, as detailed by DiSalvo in Adversarial Design [70], or Kriplean et al. who developed a platform for voters to debate contentious voting issues [137].

The most commonly used conception of political diversity is a single-dimension spectrum in order to represent a spread of viewpoints. This is commonly a continuum from liberal to conservative, often referred to as *left* and *right* (e.g. [98]) Other work uses an agree to disagree scale measured across a variety of issues (e.g. [137]). However, representing viewpoints along the a single-dimension makes it difficult to situate some political viewpoints, for example libertarianism or authoritarianism, which contain viewpoints that would make them difficult to place onto a *left* to *right* continuum. Another problem is that much of the prior work exploring political diversity relies on users making an active choice to seek out diversity in the political opinions they consume, with prior work highlighting that only a minor of people do this [97, 172].

To this end, the following study addresses both the problem of using a single-dimension of political diversity, and the reliance on users seeking out that diversity themselves. First, I use the Nolan chart [183] (see Figure 6.1), which classifies political views in two dimensions, and is thus able to represent more nuanced political viewpoints. This extends the approach taken to represent political diversity seen in much prior work. To address the reliance on users seeking out diversity themselves, I propose reconfiguring the way online news and opinions are consumed, through a smart home assistant that uses pushy characteristic to interject socio-political topics at random intervals into the home environment. Given that smart home assistants only "speak when spoken to", the VUI presents a design opportunity to interject socio-political topics into the home using a human-like voice.

I report on a study of Spkr (pronounced "speaker"), which 'spoke' opinions and views from trending Twitter discussions. To ensure exposure to a diversity of politi-

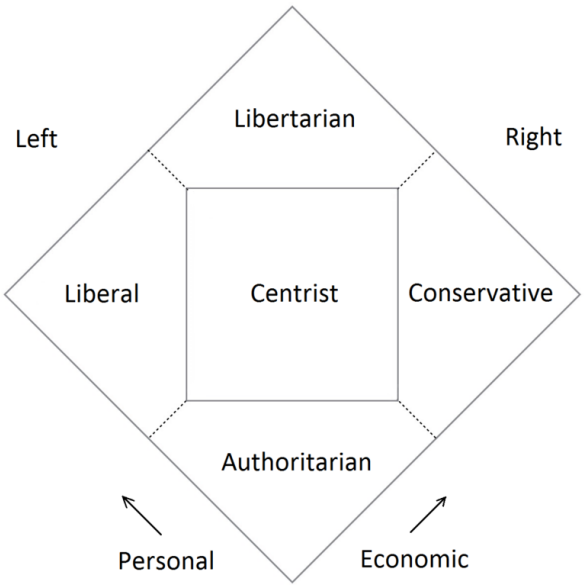

Figure 6.1: The Nolan chart of political alignment. By Proud User. Licensed under CC BY 0

cal opinions, I classified the study participants, as well as tweets from these trending topics, according to a Nolan chart. I also captured participants' responses by enabling and encouraging users to respond to Spkr. Analysis of qualitative and usage data demonstrated the effectiveness of Spkr in terms of enabling exposure to diverse political opinions and in encouraging engagement with oppositional viewpoints. The study contributes new insights to areas of HCI research tackling issues of digital democracy, highlighting the applicability of push-based design strategies in combatting selective exposure to political opinion. I further reflect on the broader ethical concerns that come from this research, and the associated challenges of implementing such systems in real-world settings in the future.

6.2.1 Political Alignment

In this study I propose the use of political alignment of both those who use the device, as well as all the tweet content played by the devices. The political alignment method is based on the Nolan chart of political alignment [183], seen in Figure 6.1. This approach provides more nuance than the common, single-dimensional spectrum of progressive vs conservative or left vs. right. The chart includes two

dimensions of rights. Social (or personal) rights include freedom of expression, personal autonomy, etc. Economic rights include the ability to own property, exchange goods and services, etc. These dimensions can vary independently. An individual may strongly support social rights, such as same-sex marriage and abortion, but also oppose granting extensive economic rights, such as free trade agreements or hands-off approaches to market regulation. Such a position would align with a traditional progressive or leftist view. Conversely, an individual may support free trade and market deregulation but oppose abortion and same-sex marriage, which aligns with a traditional conservative or rightist position. An individual who strongly supports both social and economic rights would be classified as libertarian,and an individual who does not support either would be classified as authoritarian or statist.

Numerous critiques have been levied against the Nolan chart. It does not originate in scholarly political science analysis but rather was used as a tool to explain simply and succinctly libertarian ideology [72]. The division of social and economic liberties does a poor job of accounting for situations such as prostitution [121]. Finally, David Nolan, its namesake, did not in fact originate the chart [17]. Thus, the Nolan chart may not provide a perfect analytic tool for sorting the diversity of political viewpoints. However, it does provide a useful conceptual lever that can be employed to explore possible alternatives to a single-dimensional political spectrum.

To ensure relevance to the participants, I use a version of the Nolan chart refined for the UK political context by Meek [158]. This provides a mapping of UK politics onto the Nolan chart, and a questionnaire that can be completed to align UK political views with it. I asked participants to complete this questionnaire at the start of the study, furnishing me with political alignment scores for each. Utilising these scores, I tailored content to each participant that both matches their own alignment, as well as an opposite viewpoint. I now detail the design of the Spkr device, and how I operationalised political alignment as part of the study.

6.2.2 Design of Spkr

Motivated by this previous work around filter bubbles, echo chambers and selective exposure, I designed a system that re-configured the way these discussions are presented, and subsequently engaged with. Here I build on the work of Brooker et

al. [41] who note that socio-political discussions on Twitter can become saturated with one viewpoint, drowning out other diverse viewpoints. Indeed, Doughty et al describe how such one-sided socio-political discussions on Twitter can lead to the disinhibited abuse of groups of people [74]. Twitter also presents a highly active source of discussion, with a range of viewpoints (as demonstrated in Chapters 4 and 5). In exploring novel ways to re-configure the engagement with socio-political discussion, my working assumption was that smart home technologies such as Amazon Echo or Google Home, provide a potential route to sharing such content in new ways in domestic spaces. Furthermore, their integration of VUIs (e.g., Amazon Alexa, Google Voice) provides a means for responding to such content, allowing engagement to be two way. To this end, I designed Spkr (Figure 6.2), a smart home device that is placed in the home and reads out tweets from trending Twitter discussions randomly within pre-defined intervals throughout the day. The content spoken by the device is tailored to each participant based on a political alignment assessment, resulting in the device playing a mixture of content that is based on the participant's political alignment. The result is that Spkr presents the participant with multiple viewpoints from within a trending Twitter discussion.

Operationalising Nolan Chart Political Alignment

Broadly, each day a trending discussion on Twitter would be selected, and the tweets within the discussion manually coded to identify the socio-political orientation of each message. Coded tweets are then allocated to Spkr devices, the exact mixture of content dictated by the political alignment of the participant.

Sourcing and Coding of Tweet Content Topics were selected from Twitter Trending topics, featured Twitter Moments, or by searching for hashtags or topics that were currently in the news. I maintained a running list of topics in the news that were being discussed on Twitter, with one being selected each day. I requested that members of my research group send me any Twitter topics they encountered that were being debated, so I could consider whether to add them to list. For logistical reasons, news topics were always selected and coded at a minimum of one day before participants would hear them. The criteria for selection of a topic were:

- A news item within the last 5 days
- Focus on a socio-political issue or topic
- A large volume of tweets specifically about the topic (e.g. 50+)
- Evidence of multiple viewpoints on the topic within these tweets.

Figure 6.2: The Spkr device.

Table 6.1: Example categorisation of tweets for ‘collapse of Fly BMI airline’ topic, anonymised via rewording

Category	Example Tweet
Liberal	The race to the bottom with airlines started with Ryanair, abusing passengers. Putting passengers first would be a winner.
Libertarian	Remaining in the EU couldn’t have saved FlyBMI, I don’t think. Passenger numbers were not sustainable. Political gain by blaming Brexit, maybe?
Authoritarian	Talking of Fly BMI, our local economy depends on Flybe. Big employer and connects us with parts of UK and Europe.
Conservative	Derry airport will hardly survive, following the loss of Fly BMI. When people think of NI they think of the backstop and border. But losing them will be huge for us.
Neutral / Centrist	UK Regional airline Fly BMI has announced it is filing for administration and has cancelled all flights.

Once a topic was chosen, tweets were coded to a quadrant of the Nolan chart, representing the political alignment of the views expressed within the tweet. I performed all of this coding, to ensure consistency. To do this, I developed descriptions of the stance on socio-political issues of each quadrant, derived from Meek’s model, to guide the coding. An abridged example: “Liberal - against free market capitalism, pro individual liberty". In order to code tweets, I used the quadrant descriptions and Meek’s political alignment questionnaire to assess which quadrant the tweet aligned to most. Any tweets that were difficult to align were given a best guess coding. At the end of each coding session, I consulted with Gavin Wood, a Senior Research Assistant within my research group, who helped me examine all tweets and alignments, discussing any best guess coding, where we came to a consensus.

Tweets from news organisations that did not present an overt opinion or view on the topic were used to populate the neutral category. I worked on a per-day basis, collecting all tweets for one study day before moving onto the next. An example of coded tweets can be seen in Table 6.1.

Allocating Coded Tweets The mixture of content received was based on the political alignment of the participant (see Table 6.2). For example, a participant aligned to Liberal would receive 40% Liberal, 40% Conservative and 20% Centrist content. The tweets within each quadrant represent diverse range of viewpoints, which are spoken through the device. The rationale for drawing content from the opposite quadrant of the Nolan chart was to provide an identifiable contrast to content from their own quadrant. In practice, this varied based on the tweets about a topic, and could be, using the previous example, 20% Conservative and 20% Libertarian, rather than 40% Conservative. This results from the natural variation in the viewpoints expressed on Twitter. Furthermore, we are cognisant of Munson & Resnick [44] who describe the tolerance of disagreeable news and opinion users will withstand. Therefore, Spkr provides 40% of the content aligned to their view, along with a “neutral" Centrist 20%, to maintain a degree of agreeable news. Through experimental testing, 20% Centrist (neutral) content was also necessary to orient the listener somewhat to the topic being discussed. For the purposes of this study, those participants classified as Centrist were classified to their nearest quadrant, as following experimental testing, Centrist tweets could easily be perceived as one, two, or even three of the adjacent quadrants, and thus the mixture of viewpoints was not clear.

Implementation

Spkr is composed of an Amazon Echo smart home assistant, seated on a small wooden base that contains a Raspberry Pi computer. I chose to use the Amazon Echo as it provides suitable audio functionality to play spoken content into the home. Moreover, Echo provides suitable voice recognition and conversational interface tools to support responses from participants. The Raspberry Pi uses Bluetooth to transmit audio to the Echo. The Raspberry Pi connects to a server via Wi-Fi and retrieves voice audio to be played. On the server, text-to-speech audio is stored which is generated by a voice synthesis API provided by CereProc[7]. The audio content played by each Spkr is dictated by the server, thus allowing a mixture of content

[7] https://www.cereproc.com

to be played each day. I manually input the tweet content into a database, which consisted of the tweet text, the political category, and the study day it was intended to be spoken. I wrote automated scripts that collated playlists of content for each user for each day of the study. I used a custom web interface to edit and refine voice audio, to ensure correct pronunciation. I experimentally defined the settings and voice accent used for the text-to-speech, consulting with Selina Sutton, a speech scientist PhD student, and Gavin Wood, a Senior Research Assistant, from my research lab, who helped me define a voice that would be distinct from the Alexa voice used in the Amazon Echo. This resulted in a southern Scottish-accented voice, which is geographically close to Newcastle-upon-Tyne. It was designed to use English language, which aligned to the participants native language. Prior to each piece of audio content, a short notification chime would sound, followed by a five second pause before the spoken content.

Responses to Spkr content were facilitated using an Alexa Skill, a software routine specific to Amazon Echo, which I created using an Amazon template. In order to respond to any content they heard, participants were required to say “Alexa Open Speaker Feedback", which would then prompt them with “What do you think about what you just heard?" Their following response was then recorded and stored on the Amazon servers, which I could access from the Alexa voice history interface.

Through the design of Spkr I sought to provide a sense of unpredictability, similar to the unpredictability of smartphone notifications throughout a day. Spkr was designed to operate within up to three timeslots each day, nominated by the participant, and it would play content randomly at any point within each timeslot. Following testing in my own home, and members of my research group, I decided 10 tweets per day in total would be feasible. Drawing inspiration from qualities of radio, I did not provide any replay functionality. In this way Spkr capitalises on the ephemeral nature of oral communication - that speech is physically felt by the listener, demands attention, and that the crafting of statements has an effect of how memorable they are [189].

Ethical Data Logging

The Spkr prototypes utilised hardware and software, which had been modified, or was newly developed, by myself. To understand how each participant used the device I established a set of data logging practices. In line with my methodological approach (as discussed in Section 3.3.5), I designed the data logging process to Crowcroft et al.'s principle of *privacy by design* [61], and thus collect only the required data to understand how the prototype was being used. This is especially important given Spkr would be in the living area of participants' homes, and the voice recognition of the Amazon Echo could potentially collect sensitive conversations. Primary data logging was performed through the Amazon Echo voice recognition software, which collected audio from any user addressing the Alexa voice assistant. I used the Alexa Skill to filter this data, meaning that only the participant who addressed the Alexa Skill (by saying "Alexa Open Speaker Feedback") was collected for the study. All other data addressed to the Alexa, but not directly to the Alexa Skill, was discarded. At the outset of the study I explained to participants this data collection process, and made sure to stress that the Amazon Echo would collect anything that was addressed to it (for example, if they said "Alexa what is the weather?") but that I would ignore this. Here I was mindful to make participants aware of the data collection capabilities of the Amazon Echo, so as to reduce the asymmetry between the participant, having a data collection device in their living space, and myself having access to their data [115].

6.2.3 Study Design

Spkr was evaluated as part of a four-week long study with 10 participants. The device was installed in participants' homes, playing content across 28 days. A single participant was recruited from each household. Participants were inducted into the study in their own homes, where a 20-30-minute semi-structured interview was performed, to establish existing news consumption habits, and experience with smart home technology. Participants nominated up to three time slots during the day where they were likely to be in the house and would be willing to engage with the device (see Table 6.2). Following this, a Spkr device was set up in their living space, in a place of their choosing. They were informed the device would read out opinions from

Twitter, and that they could respond if they wished. The day after the device was installed Spkr began playing content within the nominated times. This continued each day for the duration of the study. On the third day of the study, I contacted the participant to ensure the device was working correctly and that the nominated time slots were suitable, modifying them if required. At the mid-point of the study (day 14), I contacted the participants to ensure all was working ok. After 28 days, Spkr stopped playing content, and I visited participant homes to conclude the study. This took the form of packing up Spkr, and performing a 30 - 45 minute semi-structured exit interview, which explored their experiences, and the data of the topics they heard and their responses. A full debrief followed this. Participants were given £60 in high-street vouchers, broken into a £10 voucher a commencement of the study, and the remaining £50 voucher at the end of the study. One Spkr device (P9) encountered a persistent hardware issue until it was changed on day 14. I was accompanied during the entrance interviews by David Verweij, a PhD student from my research lab, and I conducted the exit interviews with partial assistance from Gavin Wood, Senior Research Assistant and Scarlett Rowland, Research Intern, from my research lab.

Participants

Ten participants were recruited for the Spkr study, who were drawn from the local population in Newcastle-upon-Tyne. Participants were sourced by a professional participant recruiter to ensure a diverse range of participants. Five identified as female, and five identified as male, with ages in a range from late 20s to late 60s (see Table 6.2). In order to maintain the smooth functioning of the study I specified the following criteria:

1. Must not have an Amazon Echo or Google Home within the house, or have ever owned one.

2. Must not be away from home for more than 5 days for duration of the study.

To accommodate for the vagaries of family life, I set 5 days as the maximum participants could be away from home. I was not expecting participants to hear everything spoken by Spkr during the course of a day, nor did I want participants to force themselves to stay at home and in the room to listen out for Spkr, as I felt this

Table 6.2: Participant information and Spkr alignment data

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Profession	Nurse	Taxi Driver	Delivery Driver	Police Office	Teacher	N/a	Admin.	Logistics	Admin.	N/a
Time slot(s)	Morning, Evening	Evening	All Day	Morning, Evening	Morning, Evening	Morning, Evening	Morning, Evening	Morning, Evening	Evening	Evening
Alignment	Auth.	Auth.	Liberal	Auth.	Liberal	Liberal	Liberal	Liberal	Liberal	Liberal
Responses per day	0.28	2.03	1.21	0	0	1.82	0.64	5.32	0.96	3.17
Responses total	8	57	34	0	0	51	18	149	27	89

would not be a natural engagement with the system. None of the participants were known to the research team, nor had any previously had an Amazon Echo in their home. All were aware of what an Amazon Echo was, with approximately half of the participants having used one at a relative or friend's house.

During recruitment there was no specification that participants should be interested in news or politics, which resulted in a diverse range of engagement with news during daily life. One participant (P9) engaged very little with news, only listening to news bulletins on the radio, whereas others (e.g. P8) avidly consumed news through smartphone apps, TV, and radio, as well as actively seeking out discussions at work and with family members. All participants are referred to by their anonymised number (Px), and researcher as R.

The topic of participants expectations of what to do with a research prototype has been discussed by Tolmie [235], who describes this as the *"propositional grammar"* of a research prototype. Therefore I was aware of priming the participants with what they should do with Spkr, and thus told them the device would read out viewpoints from Twitter, and they could respond if they wished. I wanted participants to find their own preferred way of interacting with the device, rather than implore them to respond to everything, or to think very hard about the topics they were hearing.

6.2.4 Data Analysis

All interview audio was transcribed by a professional transcription company, with de-anonymising features such as names removed from the transcript. Inductive thematic analysis was performed on the entrance and exit interview data, following the process described Braun & Clarke [37]. I was assisted by Scarlett Rowland, a Research Intern in my research lab. We both familiarised ourselves with the interview transcripts. I produced an initial codebook from an initial coding of around half of the transcripts. We then discussed the codebook, identifying duplicates and clarifying meanings. Once the codebook was agreed, I performed complete coding on the full set of transcripts. We clustered the codes into subthemes and themes, which are presented in the following sections.

6.2.5 Results

Over the course of the 28 day study, a variety of topics were presented by Spkr, predominantly UK political topics, such as Brexit, business, and crime stories. World politics also featured, such as the US-Mexico wall and the Venezuelan border crisis. In terms of the location of Spkr in participant homes, six participants had it in a lounge area (P1, 2, 3, 4, 5, 10), with the remainder having it in their kitchen.

Participants contributed their responses to Spkr with varying levels of engagement, as can be seen in Table 6.2. P6, 8 and 10 responded to a large quantity of what they heard. This took the form of short *"I agree"* or *"I don't know"* statements, to more considered responses *"Absolute fabrication, 'cos I work for a Latin American company, they pay good wages, health care, and doing right for me and my family"* [P8]. These three participants were almost always in the room at the nominated timeslot, and thus followed the content and responded to much of the content. Other participants did not respond to the device at all and had low levels of engagement. P5, a police officer, and did not feel comfortable sharing their opinion, but described that they heard a lot of the content. P3 had Spkr operate all day, and as they would be performing odd jobs, watching TV or in the garden throughout the day, and would listen and respond whenever they were nearby the device. Another group of participants (P2, 7, 9) were keen to engage with the device, but found themselves often out of the room, which was attributed to shift work (P2), or a busy lifestyle (P9). During the course of the study, the Spkr devices delivered 494 broadcasts to participants, and received 433 responses to messages across all participants. I now present the results of our thematic analysis, organised into four themes.

Spkr's effect on how news is consumed

While it was not a recruitment criteria, all but one of the participants reported being habitual consumers of news. Common news media such as television (morning or evening news broadcasts and 24 hour rolling news), news websites and smartphone news apps were used variously and habitually by participants. Opinions from social media already formed part of the news consumption process of some participants, with Facebook and Twitter being the primary platforms. Two participants (P7 and P8) stated they also used Snapchat as a source for news, consuming content from news

organisations: *"It's everything from Sky News to the Daily Mail to The Sun, it's a bit of everything under one app"* [P8]. The one exception was P4, who noted he and his family: *"Kind of live in our own little bubble, to be honest with you. We're happy we keep to ourselves, kind of thing"* [P4]. Furthermore, very few of the participants noted they would engage in sharing opinions of news stories or comment on them online, although several reported being interested in the opinions shared by others online on news stories: *"I do read* [comments] *though, always interested to see what people are saying around the issue"* [P2].

Participants noted that during the study their news consumption practices had changed or been adapted to include Spkr. For some, Spkr had become a reliable source for the news, to the extent that they did not feel it necessary to use their usual sources: *"I might have even checked the news less, to be honest, knowing that at eight o'clock I was going to get filled in with the top stories of the day"* [P2]. Others noted it had increased their interest in news: *"100% listening to more news in the car rather than music. I don't know whether it was just total coincidence or I subconsciously thought, 'There's more stuff going on out there, let's have a look and see what's happening.' Put the news channel on"* [P4]. With others it encouraged them to engage with news and discussion in ways they hadn't before: *"I've been involved in a couple of discussions at work as well, when normally I would ignore them"* [P7]. In P7's case, news stories that they considered they would not normally have been aware of were brought to their attention. Other participants presented similar views, appreciating the diversity of news stories that Spkr would bring to them: *"I think random* [news stories] *is good because I wouldn't like to say, 'I want to know this, this and this'. I think having current news, I think that's a good thing because some things you totally miss in the news."* [P9]; and *"It was pretty good what was chosen. There were a couple of things that went over my head, but overall I enjoyed listening to what came through"* [P6]. In P6's case, this often triggered a process of further research:

> *"I did look up a lot more to think I want to know myself more about that. I don't often get the chance to have things brought to light* [...] *It was good for bringing different opinions to your attention to make you question actually am I right in thinking that"* [P6].

At the same time, some participants stated they would prefer to pick the topics themselves, to include more relevant topics, for example: *"It didn't know what I liked. If I typed in my football club or 'price of diesel' or 'war in Syria' or 'Jewish history', then that would be important to me, therefore the tweets would be tailored"* [P5]. Of course, in many respects this was missing the purpose of Spkr, which was to purposely challenge the notion of being recommended news stories aligned to predisposed preferences.

Precipitating reflection upon, and criticality of, news

Many participants reported that they already employed a degree of criticality in their news consumption practices. Sometimes this was based on the perceived trustworthiness or rigour of certain sources of news: *"The Guardian, and the Independent, they've got hopefully more intelligent people with a more balanced view there, but not always."* [P2]. Others reported trying to interpret stories to identify bias, albeit quite what this was as a process was hard for them to articulate: *"I try to work things out in my head of actually what is going on. What are they telling you?* [...] *How much is it propaganda"* [P3]. However, not all of the participants exercised critical practices when consuming news: *"I never think what specific news channels will sway me to go a different way.* [...] *I'm probably really naíve to think that that's it, that's news and it's right"* [P9].

Over the course of the study, participants reported that Spkr triggered various reflective processes based on broadcast content. This included inquiring more about a story, ruminating on their own viewpoints, or triggering a discussion. P9 described how a topic about Malcolm X caused them to research the topic, and reflect on their own knowledge of history:

> *"They were saying that he was more extreme than the extremists* [...] *I had just Googled it, it said that he had been assassinated and I was just like, 'This is really quite sad that I don't know about this historical thing'"* [P9].

All of the participants were aware of contrasting or alternating viewpoints being presented within Spkr's content: *"some topics, like, they were put across and then*

I felt like the next question was put across in a different way, if that makes sense? Same topic, different way" [P8]. Others mentioned how this diversity triggered reflection on their own viewpoint: *"It was good for bringing different opinions to your attention to make you question actually am I right in thinking that"* [P6], and *"it gets you thinking about the issues so it kind of straightens your opinion out in your head a little bit"* [P2]. Being presented with diverse viewpoints led to changes in opinion for some of the participants, as discussed by P10: *"It was quite refreshing because, there was one day I found I was in agreement with it, and then they said something, and I thought, 'Yeah, you're right', and it made me give comments the other way on it"* [P10]. P9 described how the different viewpoints were persuasive: *"if it wasn't something I fully understood I would be listening to it* [...] *and then I would listen to somebody else's side of view and think, 'Oh, I never really thought of it like that'"* [P9]. Both P9 and P10 mentioned they would ask a family member about a story: *"I wasn't 100%, so on certain things I would have to ask my husband what he knew about it"* [P9].

Considering the political alignment and coding that was conducted on the content used in the study, it is interesting to note that some participants felt the content was biased on certain topics, likely reflecting the types of topics discussed on Twitter [161]: *"I think it was more left wing, yes, definitely.* [...] *like obesity, and the immigrants were mostly focused on"* [P7]. The source of the tweets was also discussed, so they could establish whether to trust it as a source: *"If I knew the background of who it had come from* [...] *I would understand, well, they've said that because of this, or they are of this background, that makes sense why they would think that way, but they are thinking that way because of that directly impacts them."* [P7]. This highlights a tension, as providing the source of a tweet could both help a listener understand the context the statement was made in, but also to introduce their own biases into this process.

Appreciating and making sense of Spkr

Overall, throughout the study participants generally responded well to the main feature and principles of Spkr. Several participants noted the ephemerality of the spoken content afforded the device with an ability to command attention: *"When that's*

actually speaking to you, then it gets your attention more, I think, because it did stop me in my track, it did make me turn around and listen to what was being said." [P7]. The lack of a repeat function lent the content a scarcity that encouraged participants not to miss it, such as rushing into the room: *"if you're through the back and you're washing dishes, or the washing machine or you're in the bathroom, it's trying to get back down in time again to catch the rest of it"* [P1]. In some cases, participants reported some frustration and upset with missing content, and rushing to the device to ensure they would not miss a broadcast as it aired: *"if you're through the back and you're washing dishes, or the washing machine or you're in the bathroom, it's trying to get back down in time again to catch the rest of it"* [P1], and *"Sometimes I was upstairs in the bath and I could hear it and I'd think, 'Oh, bugger, I've missed it'"* [P10]. In some cases participants reported turning the volume of the device as high as possible to ensure they could hear it throughout their home, and meant they avoided having to *"just sit and wait for it to go"* [P1]. While the ephemeral nature of Spkr was seen as both a strength and limitation of the device, it was clear it drove participants to engage, and to feel as though they were missing out on important material when they were unable to get to the device.

The speech-to-text voice used for Spkr also required participants to listen carefully. Several of the participant reported problems with the intelligibility of the voice, often struggling with the specific accent I had chosen: *"Then obviously being robotic as well, Scottish-robotic was a little bit tricky, a little bit thick of an accent"* [P2]. However, it was also noted how the use of a voice, especially one that required some concentration to understand, focused attention to Spkr when it was broadcasting content: *"It's easy to just put something else on the TV, but when that's actually speaking to you, then it gets your attention more, I think, because it did stop me in my track, it did make me turn around and listen to what was being said"* [P7]. Furthermore, P8 demonstrated that listening to opinions through a voice interface led to all the opinions being merged into a single persona, which they found occasionally jarring: *"Sometimes she said phrases, like she'd say something about Donald Trump and then she'd say, 'Yeah, go Donald Trump!' like that was her opinion, like she thought it was right"* [P8]. Here we see the participants attributed some kind of persona to the device, which was then confounded by the opinions that would be spoken through the device, and how it seemingly contradicted itself.

6.2.6 Discussion

In this study, I have reported on an in situ deployment of a pushy smart home device, Spkr, as a means to understand how to broaden political exposure, and to do this I used political classification to represent identifiably diverse viewpoints. I now present a short discussion of the results of Spkr, with a more comprehensive discussion being presented in Chapter 7. Here I draw out insights related to the ways devices like Spkr have the potential to broaden people's exposure to socio-political opinions and the ways they might facilitate widened awareness of and engagement with diverse viewpoints on newsworthy topics.

Broadening Exposure and Questioning Relevance

Prior research has noted that people often engage in their own critical processes when it comes to seeking news from different organisations or platforms [96]. This was also the case for many of the participants, whom were quite aware that different media outlets would spin news in certain ways, and that they might speak to distinct audiences and readerships. To a degree however letting such criticality shape choices over where to find out about and discuss news stories reduces opportunities to find out about and be confronted with alternative viewpoints on a specific issue. The participants reactions to Spkr demonstrated how this was, for most of them, a novel experience. Indeed, in some cases it is evident they were able to identify, based on the content of different broadcasts associated with similar news stories, the positions being taken and how they related to their own. In some instances hearing the diversity of perspectives made participants question *"am I right in thinking that?"* [P6]. In their work on promoting user-generated annotations of news stories, Wood et al. observed the ways that seeing news stories from multiple perspectives, and the reactions of other anonymous readers to these, users learn more about their own opinions [251]. Similarly, we can see Spkr content was used as a sounding board for one's own opinions, at times reinforcing views, but also promoting some deeper reflection on why they held a specific view on a topic.

At the same time, some participants wanted content explicitly personalised to them, as they were not interested in some topics, such as US politics, and stated they had no relevance to their life. As I've already noted in the results, the position of

these participants was at odds with the design principles of Spkr - the whole purpose of the system was to challenge individuals to have broader awareness of diverse news stories and, in particular, expose them to alternate socio-political opinions to their own. Therefore, Spkr exposes a tension between giving a user what they like, reinforcing selective exposure, and purposefully present topics that are diverse, but may not be relevant to the user. Work is ongoing with recommender systems to reduce this tension, such as Garimella [94], who propose an algorithm for exposing users to related viewpoints.

Promoting Engagement with News and Opinion

It was also clear that Spkr triggered a process of further research for many of the participants. For some this was driven by their interest in a topic. However the qualities of the system - the lack of context that often came with a broadcast, and the occasional inability to interpret a position on an issue - fed practices of seeking more details about stories. The results also show that for some participants Spkr became, at least temporarily, a regular source of news. Despite Spkr's focus on a single topic per day, several participants detailed they came to rely on it as their key source of news, as they knew it would be active at certain times of the day, such as when they got home. While this is somewhat problematic in that the single-story per day focus of Spkr may in fact reduce awareness of news overall, it presents one solution to the *news-finds-me* attitude, where users do not actively follow the news, but expect the platforms they use to expose them to all relevant and important news [106]. Whilst Spkr is focused on one topic per day, it does this in a manner to elucidate the key viewpoints around that topic.

The pushy nature of Spkr was also a powerful way to attract the attention of those around it, and to widen engagement with socio-political topics across a whole household. This was particularly effective when leveraging the household rhythm, such as during meal times, or when members of the house congregate. Spkr broadcasts were shown to initiate debate, which in some households, as P6 described, *"we never really do these days".* This points to the ability for Spkr to entangle others in the room in a conversation about topics that would not normally emerge. In one case, as detailed by P10, this even took the form of another member of the house

directing the participant what they should say, which they recounted, created some tension. Collective control of technology in the home has been explored by Baillie & Benyon [11], who describe how multiple family members use the technology, and the power dynamics that develop around ownership and control. As such, given a longer habituation to a device like Spkr, it would be possible that different family members may try to control or influence the device.

The Power of Smart Home Devices

A remarkable quality of Spkr was the power it seemed to have over the participants during the study. The ephemeral nature of each piece of spoken content was instrumental in creating a sense of urgency among those in the home, as they knew there was no way to repeat what was being said, and might miss a particular viewpoint. This afforded the device more power within the household context, as it caused participants to come running, to shush those around them, or to feel as though they had missed out on something important if they were elsewhere. Porcheron et al. [195] note how smart home assistants often enact power over a situation, for example when being invoked, and this was also the case for Spkr. As noted by Xu et al. [255], the ephemerality of Spkr content shapes the interaction of its users. Without a repeat functionality, it dictates that users listen to it, on its own schedule, even if they would not like to engage with it. As such, this study shows that pushy smart home devices are a powerful way to attract the attention, and initiate discussion with those in the household.

6.3 Summary

I have reported on the design and evaluation of a smart home device, Spkr, that used pushy characteristics to interject a diversity of opinions and viewpoints into the home. I used the Nolan chart of political alignment to provide a purposefully assorted set of viewpoints. The results show that injecting socio-political topics into a home entangled those in the room, initiating debate and discussion, where it was previously absent. The two-dimensional Nolan chart presented listeners with a diverse range of views, often outside their habitual news consumption, precipitating

further research. Therefore the study contributes an understanding of pushy smart home devices, designed with the intentions of broadening exposure to socio-political issues, can fit into the complex social household environment, demanding attention, and promoting engagement and reflection on opinions and viewpoints.

Chapter 7

Discussion

What do the studies tell us about agonistic interfaces, and what kinds of discussion, reflection and criticality do they encourage and facilitate? As a brief recap, I will summarise each of the empirical chapters, before discussing all of the results, and how these respond to the research questions. This chapter is structured as follows: First, a summary of the empirical chapters (4, 5 and 6). I will then re-present the research questions and objectives as presented in chapter 1, where I will demonstrate how and where in the thesis I have addressed them. Following this I include a short reflection on the 'implications for design' convention in HCI work, and then follow this with the aforementioned discussion and synthesis of the results. At this point it is then appropriate to include a discussion of the limitations of this work.

7.1 Summary of Empirical Chapters

Chapter 4 examined socio-political social media streams. I described the existing practices for the design of social media streams, and how socio-political content manifested. This focused on Twitter, and how it facilitates backchannels for discussion to form around TV programmes. Then, I designed and deployed a smartphone app, Moral Compass, that took this understanding of Twitter streams and reconfigured the way they are consumed. The Moral Compass re-presented socio-political Twitter streams coded by morality. I deployed the Moral Compass with 15 participants over the course of three workshops, with two homework activities in between.

Participants engaged critically with the content of the Twitter streams, and considered the audience of who would see their tags. I discussed how the compass interface was useful for discovering diverse viewpoints or as a means to filter the content participants engaged with, and how it inherently supported a diversity of views.

Chapter 5 concerned agonistic interfaces designed to facilitate criticality towards a genre of TV, reality TV, that has been identified as highly politicised, focused on socio-political issues, and is generally viewed without criticality. In order to introduce critical viewing to the reality TV genre I utilised second-screening - the act of viewing the primary TV screen whilst engaging with a second screen, such as a smartphone. I presented a discussion of existing practices around second-screening, focusing on the techniques for engaging viewers in critical viewing. In the first study I developed a smartphone app, Spotting Guide, that allowed viewers to critically reflect on the patterns and features within the programme, by tagging them in the app. I deployed the Spotting Guide in a set of three weekly workshops, with two homework tasks in between, with 12 participants. The results demonstrate that tagging is an effective way for viewers to critically engage with reality TV programmes, by allowing them to reflect on the meaning they have given to tags. Participants were also keen for a future version to introduce social elements into the second-screening process. It was also clear that such a process split attention between screens, and that this can be leveraged as a design technique. In the second study I presented the design and development of Screenr, an iteration of the Spotting Guide design. Screenr was a smartphone app for the critical co-viewing of reality TV. Wishing to understand how critical viewing can take place in situ, in the home environment, I integrated a system for the co-selection of programmes via a voting system, allowing the app to self-sustain with little researcher involvement. Screenr was deployed with 13 participants over the period of one month. The results showed that the social features of Screenr, combined with the critical viewing features, allowed different types of users to utilise the system differently, each supporting some part of the system for other users. Similarly the social aspects fostered group criticality, as well as accountability for voting choices and creation of tags.

Chapter 6 detailed the design and deployment of Spkr, a smart home device that presented an agonistic interface for consuming a diverse range of socio-political views using pushy techniques. Foremost, I analysed the existing practices around smart home devices and conversational interfaces, with a particular focus on the principle of pushy techniques. Following this I presented the design of Spkr, which utilised an Amazon Echo device paired with a Raspberry Pi. This enabled Spkr to *speak* snippets of Twitter discussion into the home. The purpose of Spkr was to present diverse viewpoints on contentious socio-political topics, and I implemented a classification system using the Nolan chart of political alignment to iterate on prior work, allowing four political alignment classifications to be used. Spkr was deployed with 10 participants for the period of 28 days. The results demonstrated that a device like Spkr can be an effective means to broaden exposure to socio-political viewpoints, and that the pushy nature of the device meant it became a regular news source, and often attracted the attention of those around it, precipitating subsequent debate. It also highlighted some ethical considerations of the future design of pushy smart home devices, which I discussed.

7.2 Addressing the Research Objectives

I will now refer back to the research objectives I outlined at the beginning of the thesis, and detail how I have addressed them.

1. *Demonstrate the motivation for challenging and reconfiguring discussion and engagement with socio-political issues.*

   Throughout the Introduction (Chapter 1) and Background (Chapter 2) I have demonstrated my motivation for this work, and have detailed how this is grounded in existing theory and prior research. I have introduced the idea of the public sphere, originally conceptualised to describe how pre-digital society would share opinions and engage in democratic debate, and from this the idea of the talkative electorate, and the various ways that citizens engage in discussion about socio-political topics, including through casual so called everyday talk. Within these communication networks and forms of communication I detailed the power dimensions enacted through modes of communication, and

how this can lead to influence and bias around certain topics, and more profoundly, the stigmatisation, exclusion and Othering of individuals and groups from society. Given this pre-digital grounding to the work, I discussed how these issues have manifested in the digital age, maintaining their existence in “off-line" society, and spreading throughout digital technologies. As part of this, I discussed how the design of digital systems can enable such powerful behaviour, and the specific problems with current designs (for example echo chambers, filter bubbles, and so on). Further motivation is presented, as I describe how a different conception of democracy, divergent from those that underpin most digital technology, such as agonism, and it's enactment through adversarial design, can address these problems. By incorporating the ideals of agonism (inclusivity, conflict), an agonistic interface presents users with a digital system that challenges the status quo, prompts reflection by users, or encourages them to engage in debate on topics that they would otherwise receive uncritically.

To further concretise this motivation, I presented a set of research questions (page 13) grounded in the existing theory and prior work, to guide the inquiry of this thesis.

2. *Understand the modes of consumption and mechanisms of discussion around socio-political topics via a thorough literature review.*

   Throughout the introduction I have presented the ways in which socio-political topics are discussed (Chapter 1). I have presented a comprehensive literature review in Chapter 2, which covers democracy and the concept of the public sphere, and breaks down how publics form in online spaces, and the ideological underpinnings of the platforms, and how this problematically affects the formation of publics and sharing of viewpoints online. As a means to address these problems, I discussed the conception of democracy known as agonism, and how this can be facilitated through adversarial design. There is also a discussion of techniques and literature that present mechanisms that could be incorporated into agonistic interfaces. Finally, the remainder of the literature

review chapter is divided into three parts, each of which discussed the particular mechanisms for consumption and engagement with socio-political topics, specific to each of the three domains.

3. *Critically reflect on existing technical practices and how this impacts engagement with socio-political issues*

   In the Methodology (Chapter 3) I introduced my critical technical practice approach, and as part of this I have included a section at the beginning of each chapter outlining the existing technical practices, and the related literature. These sections focused on:

   (a) Section 4.1 discussed the prevalence of socio-political content on social media, and detailed how this happens around hashtags. This section also focused on the nuances of the Twitter interface in prioritising recent tweets, without consideration for diversity of viewpoints.

   (b) Section 5.1 explored the practice of second-screening by TV viewers, and detailed the existing uses for second-screening apps, and the research conducted around this type of engagement. I also explained specific interaction techniques such as leaning forward and backwards that could be used by an agonistic interface to engage viewers in criticality as part of the second-screening process.

   (c) Section 6.1 discussed conversational interfaces, smart home devices, and how they could be used as an agonistic interface. This section described the existing design principles underpinning smart home devices, and the conventions for the design of VUIs that are embedded within them, such as speaking only when spoken to. I presented an argument for the use of pushy characteristics - speaking proactively - as a design opportunity.

4. *Examine the ethical issues presented when engaging human participants with potentially upsetting socio-political content, and how this affects the design of technology prototypes, and field work methodology.*

In Section 3.3 I detailed typical considerations when working with human participants, such as informed consent, the right to withdraw, and data protection. Working with potentially upsetting socio-political issues, and designing technology to encourage users to conflict is a unique context for this work, and Section 3.3 also details my analysis of these ethical considerations. This covered planning for focus groups where participants may become upset, how participants may become upset whilst using one of the digital prototypes, and how to anonymise social media data. I discussed how diversity of sample is important when a diverse range of socio-political viewpoints is required, and how participants' audio data needs to be protected on the Amazon Echo device.

5. *Use a critical technical practice approach to design, build and deploy a suite of agonistic interfaces, focused on three relevant domains.*

   I laid out my methodological approach in Chapter 3, which centred around critical technical practice (CTP), and specifically focused on the history of CTP in computing research discplines, and how it is relevant to myself as an HCI researcher, and how it can be used as a research approach. I then discussed the creation of digital prototypes as a means to study users, and the deployment of those prototypes into naturalistic settings such as the home. In Section 3.3 I discussed in-depth the potential ethical issues when studying digital technology at the intersection of HCI, CTP and socio-political topics. Adopting this methodological approach, I then presented studies around three domains:

   (a) *Reconfigure socio-politically charged social media streams to prompt reflection.*

   In Chapter 4 I presented the design and deployment of Moral Compass, a smartphone app for engaging with socio-politicised Twitter streams. Grounded in the existing practices of Twitter, Moral Compass presents users with an alternate way of viewing a Twitter stream - encoded by morality. Moral Compass also allowed users to subjectively encode how moral they thought tweets were.

   (b) *Support and facilitate critical viewing of reality TV to reveal underlying production processes.*

I conducted two related studies that facilitated critical viewing of reality TV programmes, presented in Chapter 5. In the first study (Section 5.2) I designed Spotting Guide, which used tagging of on-screen behaviours and patterns to structure criticality of the programme. In the second study (Section 5.3), I iterated upon this design by introducing sociality into the system, where I designed Screenr, a co-viewing smartphone application that facilitated group critique of reality TV programmes.

(c) *Explore exposure to diverse socio-political viewpoints through a conversational interface in the home.*

Having engaged with the existing practices around conversational agents and smart home devices, I designed a smart home device, Spkr, that used pushy characteristics to speak, unprompted, to those in the home, which I documented in Chapter 6. I used the Nolan chart of political alignment to classify socio-political discussions on Twitter, and broadcast them through Spkr into the home. Each day a mixture of socio-political viewpoints about the same topic would be spoken by the device, which precipitated discussion and reflection amongst those around it.

The final objective laid out at the beginning of this thesis was to *"Synthesise the findings from these deployments to present the strategies and techniques for the design of agonistic interfaces, and how this facilitates and guides debate, reflection and criticality around socio-political issues."* The remainder of this chapter is dedicated to responding to this objective. To guide the discussion, I focus on the three research questions: i) *In what ways can interfaces invoke agonism?* ii) *How can agonistic interfaces be designed to create and promote critical reflection and engagement with socio-political topics?* and iii) *What forms of discussion, reflection and criticality are evoked by agonistic interfaces?*

7.3 On *Implications for Design*

Before I discuss the three research questions, I feel it is important to reflect on the disciplinary conventions when discussing findings from HCI work. One commonly

used approach is to frame the findings as a set of *implications for design*, that synthesise the qualitative work into actionable system design practices. This practice of reducing rich, complex, qualitative work into simple practices has been critiqued by Paul Dourish [76]. He notes that the HCI convention of implications for design often frames the findings from ethnographic studies as a set of system requirements, that can be actioned by a software engineer. He argues that whilst it is possible to derive some actionable insights, it is important to consider the wider, more abstracted implications, or indeed that the findings may imply *not* to design. This point is encapsulated as follows:

> *"It is not that these do not have profound implications for design, because they do; indeed, often more profound than a laundry list of facts and features. Their impact, however, is frequently more diffuse. They provide us with new ways of imagining the relationship between people and technology."* [76, p.548]

I am aware that an inquiry such as this could conclude following the convention of *implications for design of agonistic interfaces*, and attempt to reduce the rich findings into a set of requirements - the implication being these could be followed as a recipe to produce an agonistic interface. Following the approach advocated by Dourish, I will produce some requirements or considerations for the design of agonistic interfaces, but I will also discuss beyond the practicable design, providing insight into the way designs evoke engagement with socio-political topics and different forms of discussion, and how this relates to wider phenomena. Given this work has been conducted with a critical technical practice approach, where throughout I have reflected upon the existing status quo of research and design in each domain, it is appropriate to discuss this reflectively. I will discuss this process, and how the process of my inquiry and it's findings can contribute to a wider understanding of agonistic interfaces, agonism and adversarial design within the disciple.

7.4 RQ1: In What Ways Can Interfaces Invoke Agonism?

The first research question guiding my thesis regarded agonism and how this can be enacted through an interface:

> *"In what ways can interfaces invoke agonism?"*

In Chapter 2 as prerequisite to explaining in what ways an interface can invoke agonism, I presented a case for why interfaces *should* invoke agonism in their design. To do this, I explained the grounding of the conception of democracy known as agonism. In particular I explained how agonism differs to traditional conceptions of democracy (e.g. liberalism), in that it is inclusive, and is predicated on disagreement between members of the public. I also demonstrated how existing inequalities stemming from the conception of democracy that underpins much social media contributes to, or exacerbates, the acknowledged problems of online publics, such as filter bubbles, echo chambers, and polarisation.

I presented adversarial design, a multi-disciplinary approach to designing digital objects that invokes agonism. Adversarial design operationalises the underpinning ideas of agonistic democracy as part of its approach to designing digital objects. It is clear from existing work that invoking agonism through an interface involves a process of identifying an existing hegemony, such as the way information is presented to users, and then examining how the interface could be reconfigured to facilitate debate or reflection. I demonstrated how previous works invoked agonism through their interfaces. We saw that these interfaces invoking agonism can take a variety of forms and functions, from the ambient radio-like devices of Gaver et al.'s Energy Babble [101] to the Newsr smartphone app that allowed users to scribble their critique directly on news articles of Wood et al. [251].

In Chapter 1, I outlined the three key tenets of agonism, and I described how agonistic interfaces speak to these tenets. In light of the empirical work presented throughout the thesis, I will now reiterate how the agonistic interfaces I have presented through my prototypes address the key tenets of agonism, and I will discuss more fully the specific mechanisms of each:

Conflict & Respect: Each of the digital prototypes presents a tangible *interface*, which can be used to engage with the fundamental principle of agonism - the respectful engagement with conflicting viewpoints. Each prototype facilitates this in a different way. For example Screenr (Chapter 5) was an agonistic interface that encouraged discussion and identification of patterns and tropes with a small group. On the other hand Spkr (Chapter 6) was an agonistic interface that brought in existing agonistic debate from Twitter and presented it to individuals/families within their home. In their own ways both of these prototypes allowed users to engage with conflicting views. In other ways the agonistic interfaces motivated users to think about their own views, and to share their views when they normally wouldn't (see section 7.5.2 for a deeper discussion of this). The agonistic interfaces also maintain mutual respect within discussion, framing the relationship between users as adversaries rather than enemies. Screenr fostered social cohesion within the group to maintain mutual respect, whereas Moral Compass made viewpoints consciously avoidable with a compass interface. All prototypes used *ad hominem* attacks as the baseline for mutual respect, and the ethical and moral implications of this stance is discussed in section 7.6.5.

Democracy & Pluralistic Debate: Agonism is founded on the principle that democracy is positive for society, and that part of a healthy democracy is debate between members of the public. Agonistic interfaces therefore uphold this principle by facilitating debate between members of the public around prescient socio-political issues, or contexts that have been identified as problematic (for example, where a lack of viewpoint diversity in a certain context has been identified as deleterious for democracy). An example of this is the echo chamber, a technological phenomenon which concerns the polarisation of online political opinions and leads to a lack of diversity in debate. Founded on agonism's imperative to facilitate debate as part of the democratic process, an agonistic interface is therefore able to puncture this echo chamber by introducing pluralistic, often conflicting and incompatible viewpoints for debate, framed in a constructive, *agonistic* rather than *antagonistic* way. The Spkr prototype (Chapter 6) was developed and deployed around this topic.

An integral part of agonism, as well as agonistic interfaces, is allowing for a plurality of viewpoints to be presented, and the prototypes presented within this thesis achieve this in different ways. The focus on plurality in the agonistic interfaces

I have described is reinforced by the absence of moderation, specifically moderation of whether a view is considered 'right' or 'wrong' in the existing hegemony. This achieves the plurality that Mouffe describes, by allowing any hegemony to be challenged, and any perspective to be expressed. In my work there is however a threshold to this plurality, necessitated by the potential use of digital technology for abusive behaviour. This threshold is partly based on respect, as I describe on the previous page, but it is also deeply nuanced, and I discuss the tensions around this issue in section 7.6.5. Agonistic interfaces specifically allow for a user to share their 'own version of events' so that they can input their perspective. For example, in the Moral Compass users were given the ability to apply their subjective morality onto a Twitter feed, and for this to be visualised.

Challenging Hegemony: Another core tenet of agonism is that a hegemony - the dominant perspective, the *status quo* - can be challenged. It acknowledges that a hegemony will exist, and this view is not excluded, but as discussed in the previous point, alternative perspectives should be included - which is dependent on the inclusion of a plurality of viewpoints. Agonistic interfaces speak to this in two ways: by challenging the hegemony through the interface itself, and challenging the hegemony through their design.

Through the interface, agonistic interfaces allow a hegemony to be challenged by allowing the continual upheaval and debate around ideas and topics. As such, there is no 'accepted view', no hegemony, that is imposed on the users of an agonistic interface, and they are designed to encourage respectful, conflicting debate between users and ideas. Of course, this is intertwined with the other tenets of agonism, particularly around plurality of viewpoints and encouraging respectful conflict. For example, Spkr (Chapter 6) encouraged a plurality of viewpoints from across a variety of political perspectives, which often contradicted or conflicted with one another. They were presented as is, and were not presented or filtered by the agonistic interface into 'right' or 'wrong', but were also presented to users without any additional contextual information about the content's author, forcing them to judge the content on its own. Another way this has been facilitated is by allowing users to explore these perspectives themselves, as in Moral Compass (Chapter 4). Furthermore, a hegemony is not forced upon the users with an agonistic interface, they are free to dismiss or engage with viewpoints as they please, but the viewpoints are always

presented should they wish to engage. However it is difficult to remove any type of hegemony, especially being cognisant of the influence programmers have on the design of digital technology, which I discuss more fully in section 7.6.5.

Throughout I have advocated that in order to design an agonistic interface the design process should engage methods that are critically informed, which work to reveal the hegemonies within a design context. Adversarial Design, as a method, is founded upon this, and the primary step is to identify a hegemony, and examine how it can be reconfigured to represent different perspective(s). By using critically informed methods such as CTP and adversarial design in this work, facilitating a plurality of viewpoints has been purposefully used to highlight the hegemony that exists around the exposure and engagement with socio-political issues in digital technology.

With this footing, the *design* of an agonistic interface challenges existing hegemonies in the given context, and purposefully (as discussed in the previous point) works to avoid hegemonies forming. As an example of this, in the prototypes I have presented, the agonistic interfaces work to address spaces where debate around socio-political issues are problematic (e.g. due to filter bubbles or echo chambers). They also work to avoid problematic polarisation and filtering within the agonistic interface itself: rather than allowing certain viewpoints to exist in isolated digital spaces, viewpoints are all presented together, therefore allowing the agonistic interfaces to actively address the plurality that is critical for agonism, and not presenting any one view as the hegemony/accepted viewpoint.

Here I have briefly mapped how the agonistic interfaces presented throughout this thesis speak to the three key aspects of agonism, and in the following section I discuss in greater detail how agonistic interfaces can be designed, and the specific mechanisms by which agonistic interfaces allow users to engage in agonism.

7.5 RQ2: The Design of Agonistic Interfaces

The second question guiding my thesis regarded the design of agonistic interfaces:

"In what ways can agonistic interfaces be designed to create and promote critical reflection and engagement with socio-political topics?"

Throughout I have drawn upon existing practices in each of the domains, and through critical technical practice have developed digital prototypes that provide an agonistic interface for engaging with socio-political issues. I have encountered numerous tensions and considerations, and in my studies propose ways to address them. I have synthesised these processes from each domain, and can provide some more generalised considerations and recommendations for the design of agonistic interfaces more broadly, that apply beyond the specific contexts studied here. However before this, I would like to discuss the value of using critical technical practice when designing agonistic interfaces, as well as discuss how consciously agonism should be considered as part of the design process.

7.5.1 Critical Technical Practice & Agonism

Before focusing on the learning in relation to specific designs of agonistic interfaces, I would like to focus on the methodology used to arrive at an agonistic interface, critical technical practice (CTP). Specifically I'd like to focus on what value CTP has added to the design process. CTP is an approach to computer science research and technology design that engages critical theory in order to examine and reflect upon existing practices of technology design, notably to examine the assumptions and biases within existing practices, as well as encouraging the re-conceptualisation of existing views towards technology design. It encourages the practice of technology design as a means to both establish and exercise this criticality, based on the idea that values and assumptions may not be easily or obviously articulated prior to engagement with practice [29].

This work is oriented around a set of research questions, derived from existing research which highlight problems or shortcomings in technology design. CTP lends itself to examining existing systems and re-imagining them using critical theory. A key part of CTP is taking a critical perspective on the design of prior technology, and examining the assumptions and biases, which may potentially be unchallenged. In practice I have used experience prototyping, which has a focus on exploring different design ideas, and communicating what it might be like for a user to engage

with a product [44, p.424-425]. Throughout the process of building the experience prototypes I have integrated my own critical analysis as part of the CTP method. Thus as a method it is focused on designing critically informed objects that respond to identified problems, and exploring how users engage with these designs.

In contrast, there are two similar methods that focus on creating designs to understand a design domain, known as *cultural probes* and *technology probes*. I will describe each of these and then compare and contrast them with my CTP approach.

Cultural Probes Cultural probes are a method introduced by Gaver et al. [99] as a means of building understanding around a design domain. Gaver et al. [99] describe the use of cultural probe 'packs' that could be given to a group of participants, which contain materials that would encourage discussion and reflection among the participants. Importantly the cultural probe materials are intended to be *completed* in some form and returned to the research team. They do not need a technology aspect, for example a map could be given to participants and annotated based on a question, or can use simple technology such as a disposable camera , which could be used by participants based on a prompt or question. It is an empathic approach, allowing the design team to understand the research population and the design domain, whereby specific designs can be produced in *response* to the experiences and issues raised by the participants through the cultural probes. It is not purely a data gathering method, as the designs are not entirely informed by the responses - thus leaving space for what Gaver et al. describe as *"conceptual interests, technological possibilities, imaginary scenarios..."* [102]). Gaver et al. suggest that cultural probes are intended to *"produce a provocative dialogue about design"* [99] and thus should be used early as a part of the design process, as a means to inform and provoke responsive designs.

An illustration of cultural probes in practice is Gaver et al.'s work around the *Energy Babble*, introduced on page 163 and discussed throughout this thesis. In order to understand the design space around energy conservation, a set of cultural probes (in the form of partially completed newspaper pages, graphs and sticker packs) were deployed with participants to understand their views on their community, energy conservation and fears for the future. The research team then used this

cultural probe material as part of their existing design process to develop the *Energy Babble* device, in effect a response to the cultural probes [101].

Technology Probes A further iteration upon the cultural probes method put forth by Gaver, technology probes (proposed by Hutchinson et al.) focus on the use of simple, flexible technologies as a means to gather understanding of a design domain, to be used early in the design process. Hutchinson et al. [123] provide a comprehensive description of technology probes: they should be functionally simple, and should be open ended and flexible to reinterpretation by participants, whereby they may use it for other, unexpected tasks. They are not oriented towards serving usability, and generally features that are undesirable to participants may/should be left in, if they are intended to prompt or challenge the participants. Hutchinson et al. describe that *"technology probes collect data about users and help them (and us) generate ideas for new technology and logging of usage data of the probes allows the researchers to discuss, with the users, these ideas"* [123, p.19]. Thus they are a means for the research team, and the participants, to build understanding and experiences of the design space. Similarly to cultural probes, technology probes are intended to be deployed into real world settings: *"A probe is an instrument that is deployed to find out about the unknown - to hopefully return with useful or interesting data."* [123, p18].

CTP, Cultural Probes & Technology Probes It is clear from this description that there is significant overlap between the CTP-informed experience prototyping that I have conducted, and the cultural and technology probes methodologies. Both serve as an exploration of a design space, and that the designs used by people are not the 'final forms' as products, but rather a provocation or reaction to a design space that leads us to understand a specific context or problem.

It is evident that CTP and cultural/technology probes share common traits. For example the focus on openness to subjectivity by participants, and the emphasis on learning through real world use. Gaver places emphasis on the subjective nature of cultural probes: *"we use this purposely uncontrolled and uncontrollable approach to help us understand design domains in new ways"* [102]. My approach shares

this openness to subjective individual feedback, and as it is critically-informed by the traditions of computer science research, I have purposefully incorporated this given (as discussed earlier in section 7.3) the tendency for engineering/computer science oriented work to focus upon data gathering as a source for system requirements, ignoring the qualitative richness.

Hutchinson et al. note that a technology probe *"is not a prototype, but a tool to help determine which kinds of technologies would be interesting to design in the future"* [123, p18]. For their Energy Babble work, Gaver et al. used cultural probes as a means to understand their design space, and built the Energy Babble as a response to the cultural probes, as well as their existing design conceptions. In contrast to this, my approach involved critical examination of the medium (technology) and the related assumptions and biases around the design of it, and conscious consideration of agonism to create a digital prototype that addresses this design space. Both approaches resulted in an agonistic interface (as discussed on the following pages), but use a different method to map the design space.

A further similarity is that technology probes do not serve usability, in the sense that if a probe is intended to challenge a user, it does not get redesigned based on the users feedback that it is challenging. This is similar to my approach, where aspects of the technology prototypes deliberately present a challenge to users, however these challenges are purposefully introduced during the design/development process to challenge a behaviour/design pattern identified during the CTP process.

A crucial difference between the two approaches is how ideas about the design space are assembled. In my work, my understanding of the design space is assembled through the use of critical technical practice, which draws upon previous work to identify acknowledged problematic designs, and the practice of exploratorily designing experience prototypes to populate the understanding of the design space. In contrast, when using cultural/technology probe methodology, the knowledge of the design space is built through the deployment the probes themselves, which are designed purposefully to evoke open ended responses, the use of which helps build knowledge of the design space, which is subsequently used to inform further designs. Both CTP and probes produce insight into the future design space, but CTP front-loads the critical analysis of the design space during the design and development practice. Thus my approach focuses on how prototypes grounded in a criti-

cal analysis of existing technology are used by participants, and cultural/technology probes use probes as a means to build an understanding of a design space through deployment of the probes themselves.

Cultural/technology probes are also used as a means for the research population to interact with the design team (as in [101, 235]), where bi-directional communication may occur, expanding the knowledge of the design team and informing iterations of their design responses. In contrast, my critical technical practice approach does not engage in such bi-directional communication. I have demonstrated how I have addressed an existing problematic context, using critical theory to inform the redesign of technologies into prototypes, with the interaction between participants and the prototypes occurring at the end of this process.

In describing technology probes, Hutchinson et al. note that technology probes help meet *"... the design goal of inspiring users and designers to think of new kinds of technology to support their needs and desire"* [123, p18]. As such probes are used as tools to aid understanding, whereas the prototypes presented here represent designs grounded in existing contexts and the redesign of them to address identified problems. The use of critical theory during the practice of developing the experience prototypes (CTP) grounds them in existing technology and existing problems, and thus the resultant experience prototypes propose, through their design, a potential solution/challenge to identified problems. This orients them away from exploratory cultural/technology probes - which may be purely provocative (in order to glean understanding from participants), and may be discarded after use in the study. Therefore the four experience prototypes I have presented are oriented towards prototypes, which may be iterated upon, as they are derived from re-configuring and re-conceptualising existing approaches in light of critical theory.

Consideration of Agonism As noted in the introduction (page 9) my use of the term *agonistic interfaces* is a useful articulation of both tangible interfaces that allow users to engage in agonism, and interfaces which are the product of agonistically-informed design approaches. In each of these cases, the end product is an interface through which to perform agonism, but the process of arriving at this interface - through a design process - can differ in terms of how consciously agonism is considered throughout the process.

Two examples discussed prominently throughout this work are the Energy Babble [101] and The Social Printers [107]. As discussed in Chapter 2 both of these projects speak to some of the tenets of agonism, as they are focused on linking publics together and encouraging engagement with socio-political issues. However their designs were not grounded in adversarial design, but rather an examination of a design space via cultural/technology probes. Whilst their approach is not agonistically-informed, their design process nonetheless led to designed objects that speak to the core tenets of agonism (as discussed earlier in this chapter), and thus I would describe them as agonistic interfaces.

In my work I have taken similar focus on socio-political issues, but by using the agonistically-informed adversarial design approach, my work has explicitly oriented the designs around the aspects of agonism (discussed earlier in this section), and thus there is an overt and conscious consideration of agonism in the process, which results in agonistic interfaces.

Therefore it is possible to arrive at an agonistic interface by consciously considering the principles of agonism, as well as by engaging with a broader design space, which leads towards the principles of agonism. As a counterpoint, and in order to demonstrate this distinction, approaches that are not specifically agonistically-informed, such as critical design (where the principles of agonism *may* be part of the design processes, or may also be absent) can lead to designs that engage the principles of agonism, but they may equally lead to designs that are purely provocative, and do not allow for users to engage in agonism at all.

Therefore, to produce an agonistic interface it is not a requirement to consciously consider agonism in the design process, but by engaging agonistically-informed methods it ensures that designs will consistently speak to many, or all, aspects of agonism, and I argue that it is an important consideration when setting out to design an agonistic interface.

7.5.2 Difficulty & Friction

It is clear that asking that TV viewers or users of a system engage in *active* criticality, where they must input considered data, or watch closely for something, creates difficulties for the user. This ranged from being overwhelmed at the amount of data

being produced by others on Twitter (Moral Compass), and struggling to pay attention to two information rich screens (Spotting Guide & Screenr), to being interrupted in their home and rushing to listen to a device speaking (Spkr). Throughout I have argued that these mechanisms, whilst challenging for the user, are necessary and useful in order to prompt engagement with socio-political issues.

In Section 5.2.3 on page 109 I discussed the idea of “Designing for friction" as an outcome of the Spotting Guide study, where I drew upon Korn and Voida’s idea of designs for friction, designs that *“do not want to help you; rather they place little obstacles in your way"* [135, p.8]. The act of spotting patterns and themes in the programme was not an easy interaction for the viewers, as they were splitting their attention between both screens, and this is something they recounted in earnest. However the process encouraged viewers to engage more critically than normal viewing, as they noted they were drawn into the activity, paying more attention to the content of the programme. This echoes Chapter 4 where Moral Compass users experienced a point of friction as they were put into an interpretative position with little context for the tweets they coded, encouraging them to reflect on their own position to the tweet they were being asked to code. Here we can see that integrating frictionful activities into the design of the agonistic interfaces *prompted* the reflective and critical processes around the activities, which I argue would have been less effective had the activities been seamless and easy to use for participants.

There is, by design, a tension here between entertainment and engagement in the second-screening context. TV programmes are intended primarily for entertainment, but applying a task that requires concentrated engagement is at odds with this process, as was demonstrated by the participants. The solution is presented in balancing between intense activities, such as spotting or tagging, and less intense activities, such as reflection, and even completely disengaging. This was bourne out across the studies, and in particular in the second-screening work (Chapter 5), where participants evidently discussed their use of time to “lean back" and disengage from the criticality, due to the difficult and intense nature of critical second-screening (see page 134 for my discussion of this). Similarly, Spkr’s participants knew there was a finite amount of content per day, in specified time slots, so they could relax outside these times. This shows us that as part of an experience with an

agonistic interface, there should be periods of activity, and period where users can switch off their critical processes.

7.5.3 Attracting Focus

In a specific context, it was evident that some modes of interaction were powerful, in that they prompted users to engage or pay close attention to the digital prototype. This section is dedicated to discussing these powerful interactions, and how these design decisions worked to engage users:

Liveness and Presence An understanding of the medium (as advocated by DiSalvo when discusing advesarial design [70]) has allowed me to design interfaces that take advantage of features of the medium. As outlined in Chapter 5, TV viewing is often conducted with others physically present in the same room, watching the same programme. A motivation for developing Spotting Guide and Screenr was to introduce the user to external critique from others who were watching the programme, but not necessarily present. When interviewing participants it was clear that being engaged with criticality through their smartphone allowed the critiques to feed back into the room, and engage the others they were watching TV with. This is an established practice, as highlighted by Anstead et al. who detailed how those in the same room would share video through their mobile devices when second-screening a sporting event [8]. Liveness, being present simultaneously with others, thus presents an opportunity for engaging users in agonistic debate, and can be considered differently across contexts beyond live TV. In a previous study exploring online activism and their use of digital technology, I examined how activists used liveness as part of their technique for disseminating their activist message [84]. This was achieved by inserting their message into a Twitter stream associated to a TV programme when it was live on air, which meant many people would be looking at and participating in the Twitter stream, and thus see and engage with the message.

Along similar lines, presence in the physical environment is also a useful way to engage users. As demonstrated in the Spkr study (Chapter 6), those in the same room were often, willingly or unwillingly, involved in the process of comprehending, understanding or reflecting upon the content being spoken by Spkr. This was also

the case for Gorkovenko et al.'s Social Printers, which engaged all in the room who were watching TV [107]. In the context of the home, there is often a rhythm and routine to the movements of members of the household, who congregate in different spaces for different purposes. Therefore the household rhythm is a rich, often unique situation in each home for a device to operate within [30, 148]. Dickinson et al. [68] demonstrate how external factors such as employment and family arrangements largely define the household rhythm, and influence the way media is consumed within the house. Predicting an appropriate time, within each unique rhythm, for interruption from a pushy device could be addressed in future work, e.g., using machine learning, engaging the household in discussion using a VUI, or using externally scheduled events such as political events. However I maintain that the unpredictable nature of the interruption, even within a specified time window such as with the Spkr deployment, encouraged participants to wait and anticipate when the device would speak, and lent the content some urgency, as they would try not to miss the content, as they could not predict it.

This idea could easily be extended to other contexts outside the home where multiple people are present, but consideration would need to be made for the different tensions and challenges presented, for example by a workplace, or a pub [196], which have their own kinds of social rules, conventions, modes of speech and interactions, which should be accounted for.

Leverage Social Responsibility Social factors were a useful technique from a practical perspective, in order to organise and sustain co-criticality. Screenr was designed purposefully to use open voting, with the voting choices visible to all users, to invoke social responsibility in the voting process. This resulted in users discussing their voting choices before casting the vote, and feeling responsible for what was eventually viewed, or what they voted for. Here participants wanted to vote for the winning show, or would reprimand their peers for a poor voting choice. Surfacing the voting choices of the group each week also encouraged users to vote for a mutual option. Whilst this group dynamic may be useful in coordinating group co-viewing, there remains the continued pressure of individual schedules. The co-selection through open voting mechanism of Screenr presents a means to reconcile these pressures and achieve a useful, yet participant empowering, experience.

Ephemeral interactions Another powerful mechanism that was useful for attracting focus is *ephemerality* within the user interface, which Doring et al. define as:

> *"Ephemeral user interfaces are a class of user interfaces that contain at least one UI element that is intentionally created to last for a limited time only.* [...] *ephemeral user interfaces provide a rich and multisensory user experience. They may deliberately be designed to offer only partial or imperfect user control."* [73, p.77]

The Spkr device was designed purposefully without a repeat functionality, and thus the spoken content was ephemeral, as it would only be spoken once. Participants knew this, and understood that what they missed formed part of the topical discussion for that day, and which caused some participants to rush into the room, trying not to miss what was being said (as discussed earlier in this section). The Screenr interface used ephemerality when displaying other users' tags, where all tags were shown in chronological order, with older tags disappearing from the screen, and could not be used by other users thereafter. Participants explained this made them engage with the tagging interface more, driven by their curiosity of what other people were saying, and that they knew the tags would disappear if they didn't look.

Integrating ephemerality into an agonistic interface therefore encourages participation in the agonistic debate. Rather than allowing them to participate whenever they would like to, they are required to actively wait, or listen, to their opportunity to engage with discussion. I would like to stress here that a balance must be achieved between ephemerality and permanence. As was demonstrated by the Spkr study, whilst the ephemeral content can encourage users to engage with the system, as part of their life they are also balancing other pressures, temporal, environmental, etc, and it is possible they can become disinterested or simply unable to engage, if they are not able to engage with some degree of convenience. Here I advocate this balance can be achieved by assessing the environment - as with Spkr, narrowing the times when the device *might* speak allowed participants to predict when the content would play, and integrate it into their household rhythm.

Pushy characteristics In Chapter 6 I described my rationale for using pushy characteristics as part of a smart home device. I will briefly revisit this discussion, but the full discussion can be seen in sections 6.1 and 6.2.6. Based on the Spkr study, it is clear that using pushy characteristics lends the device a number of attributes that would not be afforded otherwise. Smart home devices generally speak only when spoken to, and thus by enabling the device to speak unprompted it allows the designer to invert the relationship between those in the home and the device, in that the device takes on some agency of its own. In the Spkr study, this meant that participants came running to hear what was being said, as it was not always speaking at a convenient time. This led to supportive practices in the household, such as other family members relaying what the device had been saying. Whilst to some degree this was to support the study (the other household members were helping the participant to respond with data for the study), it also worked to precipitate discussions within the house, as those around the device engaged with the content.

This mechanism has also been analogously explored in previous work, which has demonstrated how the passive, or unintended, engagement with a device or content is effective at prompting reflection or debate. We can see that the Datacatcher device, and Energy Babble device, both of Gaver et al. ([100, 101, respectively]) embedded themselves into the participants' daily lives, and as such the information that was being produced by the devices was cause for reflection by the participants, or it was discussed further with those around them. Similarly, Gorkovenko et al.'s Social Printers were designed specifically to encourage those watching TV together to reflect and discuss on the topics and questions being printed out [107].

There is a clear desire for smart home devices that use pushy characteristics: Beneteau et al. have recently documented that parents are using their Amazon Echo devices to develop their own pushy characteristics within the home environment. They explain how parents used timers to announce to their children it was time for bed, and they describe how the somewhat neutral status of Alexa within the family lent it some authority [18]. Furthermore, Bowden et al. showed that some users want their VUI to be able to contribute things itself, have it's own stories or opinions, to give the device more of a personality and allow them to build rapport with it [31].

However I would like to offer a note of caution here. It is easy to see how such a technique could result in detrimental experiences, and deeply unethical practices,

through certain types of use. Indeed, as Kirman et al. [133] highlight, pushy devices can be used to enforce, or reinforce, specific behaviours and what some may consider 'desirable' attitudes within the home. In applications dealing with socio-political content, those who control the design and functionality of smart home device (e.g. the developers) are in a position of power to manipulate the framing or sourcing of the content being drawn upon. This power would facilitate explicit and implicit bias, as is common in agenda practices of news editors [155], and given the sensitivity of some socio-political topics, could severely impact the democratic process or further the stigmatisation of vulnerable groups. Given the potential for different voice user interfaces to be seen as more trusting than others based on their vocal qualities [228], voice interfaces like Spkr could be seen as a route to broadcasting specific political views into people's homes and shaping householders socio-political opinions. Furthermore, through requesting responses from households to news content, devices like Spkr could feasibly become 'political sensors', that could be used as a means to gather political opinion. As such, while the intention was to use Spkr as a provocation for participants, and with positive intentions around promoting wider engagement with socio-political opinion, it is easy to see how technologies like this could be used nefariously to do the exact opposite. Concern is already being raised about the agency of smart home assistants (and virtual agents more generally), as discussed by George [104] who asks whether they should have an emotional capacity. For example, *"If an illegal but safe private action is detected by home system sensors, what is the agent's permitted action (inform the system owner; call the police, realise there is no danger and protect privacy), and who determines this response setting?"* [104]. Therefore it is clear pushy characteristics represent a powerful interaction that needs careful consideration before they are used.

7.5.4 Facilitating Diversity

A fundamental requirement for an agonistic interface is the ability to facilitate a diversity of viewpoints - in order to do the work of agonism an agonistic interface must be *inclusive* of multiple viewpoints. I have demonstrated in the studies how this can be achieved in different domains through the design of the interface.

In the second-screening context, Spotting Guide allowed users to create their own tags to describe the behaviour they were seeing on screen, and even though they were provided with some pre-populated tags, they were able to assign their own meaning to them. Iterating on this design, Screenr introduced social features to the tagging process, allowing for the diverse interpretations of the programme to be shared within a group. Furthermore, tags were not shared with any description of what they meant to the author, and thus could undergo a process of appropriation by other users, who could build their own set of tags that represented what they saw in the programme. However, they still engaged with the tags and messages produced by other users, and were exposed to others' views throughout their use of Screenr.

As discussed in Section 5.3.5, creating a range of ways for users to exercise their criticality, whether through production of tags, spotting tags within the programme, discussion with others, or a mixture, allowed users to fit into preferred social roles. Thom-Santelli et al. describe how such social roles emerge within social tagging systems, with a producer-consumer relationship evolving [232] - something which was evident in the Screenr study. Here I advocate that future agonistic interfaces can harness this producer-consumer relationship to benefit the critical process. By allowing some users, who may be more keen to produce critique, to input their critique into the system, it allows others, who may be more inclined towards reflection, to consume their critique. It is important however to allow all users of such an agonistic interface to fulfil either role and fluidly switch between them, in order that their viewpoint is included, should they wish to share it.

The Moral Compass (Chapter 4) also facilitated diversity through its design by allowing the Tagger (who was assigning morality to the tweets) to share their views, and did not feature any pre-created moral codes, providing them with a blank interface that they could populate themselves. During the deployment and interviews with participants it became clear that the compass interface could be used as a filtering mechanism, allowing users to view their preferred moral points of the Twitter stream. One concern raised by a participant was that such an interface could facilitate censorship, given that users could avoid tweets with a certain morality. A future design could actually leverage this type of usage, for example, by prompting users to view moral points on the compass that they haven't viewed: "you haven't viewed this compass point yet, is there a reason?", or "are you aware that [opposite viewpoint]

is being said over here?" This would help to re-cast the idea of censorship as a resource for critical thinking and provocation - i.e. if you are only viewing the content you want, you might unthinkingly be censoring an oppositional view. Being urged to reflect on this filtering, rather than simply being presented with filtered information may prompt deeper thinking about why you did not want to see it.

7.5.5 Discoverability and Exposure to Content

Agonistic interfaces should ensure they expose users to diverse content, and this can be achieved by allowing users to discover it themselves, through the interface.

Referring back to the discussion of personalisation and agonistic interfaces (in Chapter 6 on page 159), as users we are used to personalisation algorithms as part of the platforms and systems that we use, and thus when they are absent, or as in the case of the Spkr study, purposefully breeched, it highlights a tension between what some users may consider a desirable user experience and the idea of broadening exposure to topics and issues. As Garimella et al. [94] demonstrate, research is ongoing to explore the best way to introduce external and conflicting views via recommender systems.

In order to do the work of agonism, it is therefore important that an agonistic interface nevertheless present a diversity of viewpoints to users, whether they would like to see it or not. As has been demonstrated by the Moral Compass (Chapter 4), Spotting Guide, Screenr (Chapter 5) and Spkr (Chapter 6), users may chose *not* to interact or engage with a provocative viewpoint, but they have still experienced or seen the viewpoint. I argue that it is important for users to experience diversity, whether they find this a good user experience or not. This work is motivated by problems with the way socio-political topics exist in online publics, and I include a quote of Carl Miller, which highlights the need to reveal and confront these problems:

> *“Algorithms are sometimes too complex for even their creators to understand. Protocols enforce rules too arcane for us to pay attention to. Bots are just one example of automation doing things that previously only humans did. Each example is different, but each shows how the basic way that our lives are covered by a kind of camouflage that we can’t, and don’t, look under."* [164, 310]

I am advocating that by introducing viewpoint diversity to users - whether they want it or not - works to lift this camouflage and reveal the manipulations of TV producers (Chapter 5), the filtration and segmentation users of social media are subject to (Chapter 4), and the broad range of socio-political opinions in contentious debates (Chapter 6).

The mechanisms I have demonstrated through the design of the digital prototypes present further opportunities to encourage users to discover viewpoints themselves. Pariser describes how filter bubbles threaten to remove serendipity from the news consumption process [191], and here I posit an agonisitic interface such as Spkr reintroduces a degree of serendipity into the news consumption process, by pushing socio-political topics, that may be minor news items in mainstream news outlets, but are being actively discussed online. Serendipity presents opportunities for the design of systems when considered as a value in itself, even if this is challenging for the users [132]. Moral Compass and Spkr both harness serendipity, as they present topics to users without any input of their own, which may pique their interest, or force them to reflection on an issue.

7.6 RQ3: Discussion, Reflection and Criticality through Agonistic Interfaces

The third research question guiding this work concerned discussion of socio-political issues, and what effects agonistic interfaces may have on, or what we can learn about, the way people debate and discuss these issues.

> *“What forms of discussion, reflection and criticality are evoked by agonistic interfaces?"*

7.6.1 The Concentration of Discourse

In Sections 2.1 and 2.2 I described the idea that the public sphere does not translate directly into the online space, as Habermas himself noted, as publics are fragmented, and often shaped by commercial motivation.

Thus we can see that an agonistic interface operationalises the *concentration* of discourses, as per Gaver et al. [101], who proposed concentration as a third tactic in DiSalvo's framework for the construction of publics [69]. This happens by bringing together discourses that are often conflicting. Gaver et al's Energy Babble brought together discourses around energy conservation, whereas Spkr brought together perspectives on socio-political topics, that would not often be consumed together. Similarly the Moral Compass reconfigured the way tweets were interpreted by adding a new layer - the morality of a tweet - around which tweets were grouped.

Concentrated discourses can also become a reliable source for news, as I saw during the Spkr study. Once participants understood that a variety of viewpoints on the same topic would be presented at a set time each day, some of them noted they actively viewed the news less, as they knew they could rely on the device to give them something to think about. This ties into existing modes of consumption for news topics, such as the incidental exposure to news topics that comes from engaging in a leisure activity (e.g. browsing social media for personal stories [27]), and the *news-finds-me* principle, as discussed in Section 6.2.6 on page 160.

I have already discussed in Section 7.5.3 on page 185 how devices and systems that concentrate discussions about socio-political topics present the possibility for manipulation and bias, which links back to Habermas' point about the online public sphere (see Section 2.2), that they are underpinned and influenced by commercial interests. Throughout this thesis my digital prototypes have been developed and deployed as part of research studies, with the data carefully controlled. When developed as part of a commercial system, it would be possible to monitor involvement in agonistic debate, or even use it to inform the government. I would also posit that an agonistic interface sits on a foundation of a belief in democracy, and thus is not easily compatible with a non-democratic society. The privacy and security issues around smart home devices have already been raised, given their data is being used as evidence in court in some countries [199]. The concentration of socio-political discussion into a few/one device therefore needs careful regulation, the beginnings of which are presented by the growing privacy initiatives around the protection of digital data, spearheaded most recently by GDPR [202].

7.6.2 Creating a Space for Reflective Inquiry

A unifying feature of all the agonistic interfaces presented is that they provide space for reflective *inquiry* (see page 29 for a discussion of Baumer's principles of reflection in HCI). As I will explain, the digital prototypes facilitated reflective inquiry in different ways.

Providing a social space for reflective inquiry, bereft of expert opinions guiding discussion allowed users to discuss topics in their own terms. Many participants in the Spotting Guide and Screenr study described their existing second-screening behaviour, wherein they would monitor social media but often not post content themselves. When comparing those forms of engagement with the behaviours shown using Spotting Guide and Screenr, it is clear that it provides a relatively *safe space* [249] for people to air their views, supported by the small group setting. The chat feature of Screenr allowed for a stream of commentary, enabling quick back and forth comments which provided an element of social cohesion. Tags were less social as they were anonymous, allowing users to share in a “fire and forget" way. These elements carry highly different social implications and we can see some of the disadvantages of social media borne out in Screenr, including some flaming activities, but the examples here were modest and were largely offset by expressions of group cohesion. It was interesting to note the willingness of participants to provide a social context for their viewing and I would suggest design elements that support the inclusion of other members of the family, in the same room, might be worth further consideration.

Some of the mechanisms used within the agonistic interfaces prompted wide ranging reflection on socio-political issues, just by using the agonistic interface. The act of reducing on-screen patterns into short pithy tags as part of Spotting Guide caused some participants to reflect on the issues they were having to reduce into tags. Notably one participant described how this led them to reflect on the difference between “drugs" and “cigarettes", reconciling how they were using the tags, but also why they, and by extension society, differentiate the two. Similarly participants also described how Spkr became a sounding board for their own opinions, where they would abstain from responding to the device if they were unsure of the topic, waiting to listen to other perspectives (see page 156). In some instances this also

led to participants researching the topics themselves, through their usual process when researching news. Thus a device such as Spkr not only works to broaden exposure to socio-political issues, but can also act as a cross-cutting source for news discourse, that is not directly a news outlet but informs the users opinion, such as a political blogs [114] or fan forums [108].

We can see here that such agonistic interfaces can either present a *tangible* space for reflective inquiry, such as a chat or tagging system, or a *temporal* space for reflective inquiry, such as piquing a user's interest or challenging their viewpoint at an opportune moment.

7.6.3 Self-Censure & Missing Context

An interesting effect on users of the agonistic interfaces was that it highlighted their own practices of self-censure, and their desire for context when they were engaging criticality.

As I have noted on page 111 in section 5.2.3, when confronted with a wide range of different views and opinions users often find it difficult to identify where they should situate their views. This shows that providing a mechanism to tag or encode a meta analysis of pieces of content (e.g. on-screen behaviours in Spotting Guide, the morality of a tweet in Moral Compass) prompts criticality on the socio-political topics that underpin it. By introducing social aspects to the agonistic interfaces - that their tags or encoding would be shared and viewed by others - activated participants' desire to self-censor, to reflect on how their views fit with others, and how or if their views would be challenging to others. As such, users of agonistic interfaces in this way are thinking of their imagined audience [145], by reflecting on how their views would be received. I stress that whilst self-censure could be considered as actually reducing diversity, as I have explained earlier in this section, by concentrating discourse and providing supportive spaces through an agonistic interface, users are facilitated to share openly, even though they will still consider the reception of their views.

Many of the agonistic interfaces I deployed forced users to exercise judgements on data with limited or reduced context. Opinions spoken through Spkr were stripped

of identifying features, so it was not possible to know who had posted the tweet. Similarly tags in Screenr were anonymous, and tweets were also stripped of identifying features in the Moral Compass Tagger. In all of these studies participants discussed how this lack of context made the critique and reflection processes difficult, as they were not able to make judgements based on the meta-data around what they were engaging with. In removing the extra contextual information (such as the Twitter account that produced the tweet), this frustrates the categorisation process, as such information is used to determine the motivation of the author, and what angle they are approaching a topic from. This information is also used to appraise the author as to whether they are a credible source, as part of the Othering process. In stripping this information, users are required to categorise a tweet (for example) purely on it's content, rather than an appraisal of the author, and the author's relation to society. It is also an example of friction as part of the design process (discussed in Section 7.5). The agonistic interface is therefore actualising inclusive debate, ensuring that debate is conducted in an inclusive way, and presented via a medium where all participants are respected, but their views can be resisted equally, as part of agonistic debate.

7.6.4 Revealing and Precipitating the Socio-Political

The digital prototypes provided an insight into the way discussions of a socio-political nature were precipitated between users and those around them, as well as how the users did some of this work - revealing socio-political issues through criticality - themselves.

Bringing users together through the social aspect of Screenr allowed for group criticality to build. As I touched upon in the previous section, the absence of any TV viewing expert, or of the researcher (myself) within the chat discussion meant that discussion tended to be free form, where they would approach criticality in their own terms, relating to their own collective experiences and viewpoints. Whilst the agonistic interface provided the space and the context (second-screening) for their criticality, it was the participants themselves who brought their own critical practice into the group, and reconciled their views with that of others. The act of inserting

opinions about socio-political topics into the home environment, through Spkr, created debate amongst family members according to the participants. With the device located in shared family areas, such as the kitchen, participants described how it would talk during family dinner time, often starting them off on a discussion about what had been said. Given that each broadcast from Spkr was a single opinion, this would still be enough to create a debate between the family members.

The agonistic interfaces provided a structure that allowed users to direct their criticality towards a specific context, for example reality TV, a genre highlighted as being problematic with respect to the representation of those featured in the programme (see Section 2.4.3 for the previous work highlighting this). In the Screenr study it was clear that criticality could be exercised onto a variety of reality TV programmes, allowing the viewers to build their own critique of the programme each week. This goes some way to address the issues presented by the genre, despite programme makers having repeatedly defended their productions as surfacing socio-political issues and highlighting these to the general public [193]. Whilst not unpacking the underlying issues behind the programmes, Screenr does create, through social interaction, the means to surface these wider issues. Such critical collectives may be easily formed through fan communities for reality TV (see [216] for a discussion of the fan communities around The Only Way is Essex and other reality TV programmes), where Screenr could be used as a means to discuss and critique the programme's content.

7.6.5 Morality and Acceptability

The idea of agonistic plurality is such that all those who wish to share their views are able to do so. The purpose of agonistic democracy is not to reach a consensus, but to continually discuss, through challenge and conflict, the collective issues. It is a politics of inclusion, where all views are welcome, and are treated with respect, but it is founded on the understanding that people will not agree. However, it is clear also that online communication technology has facilitated ad hominem attacks, flaming, bullying and abuse (e.g. [149, 194]), which technology designers have tried to control and limit. It has also been a site for the sharing of conspiracy theories, alternate histories and fake news (e.g. [5]). Whilst platforms have developed robust

strategies for tackling personal attacks, they have historically been slow or unwilling to act around conspiracy theories and fake news [166, 203].

Given agonism's imperative focus on conflict, it is therefore important to discuss this dichotomy, between accommodating all viewpoints, and how this is reconciled with extreme viewpoints or abusive behaviour. During my studies I did not encounter any behaviour or viewpoints that would be considered abusive, or mis-informative, (likely due to participants being cognisant they were in a study) but there was potential for such behaviour or viewpoints to arise through the use of systems that encourage open, plural debate between individuals - a risk that I discussed in my Ethical Considerations in Section 3.3.

Thus we are presented with a question - 'how far' should agonistic interfaces go? Should they accommodate what most would consider extreme viewpoints? They are indeed part of the plurality of viewpoints and thus we shouldn't exclude them - but how is that reconciled with cultural and societal expectations of what constitutes democratic debate (even with conflicts), and what represents abusive or unwarranted topics for discussion. Mouffe tackles this topic directly:

> *"No doubt, respect is necessary among the adversaries involved in an agonistic struggle , but one important question needs to be raised concerning the limits of agonistic respect. Can all antagonisms be transformed into agonisms, and all positions be accepted as legitimate and accommodated within the agonistic struggle? Or are there demands that need to be excluded because they cannot be part of the conflictual consensus that provides the symbolic space in which the opponents recognize themselves as legitimate adversaries?"* [169, p.13-14].

As we can see here, Mouffe is reinforcing that agonistic debate must involve respect, and thus systems and platforms are justified in their efforts to moderate dis-respectful behaviours such as bullying and ad hominem attacks.

Mouffe goes on to argue that in order for an agonistic struggle to be truly agonistic, citizens must be able to challenge the existing hegemony, i.e. the cultural rules set by those in power. In excluding topics for legitimate debate, it immediately creates an *us vs. them* relationship, effectively Othering a set of views, and establishing a hegemonic view. Thus, truly agonistic debate should always be open and

receptive to any topics to be broached as part of the respective debate between legitimate adversaries.

To illustrate, Mouffe presents the principle of human rights. This is part of the hegemony, it is considered as universal in modern (global) society, and should not be challenged. However, to challenge the idea of human rights, in an agonistic sense, should be possible, as it does not dismiss the idea of human rights itself, but rather challenges the interpretation of human rights as it exists.

> *"What Western culture calls 'human rights' is in fact a culturally specific form of asserting the dignity of the person, and it would be very presumptuous to declare it to be the only legitimate one."* [169, p.31].

Thus other interpretations of human rights may exist, that conceptualise dignity in a different way. It would therefore not seem unreasonable during an agonistic debate for discussions to cover the variety of conceptions of human rights, and thus challenge the existing Universal Declaration of Human Rights.

It is therefore provocative to consider the effect of an agonistic interface, used beyond the controlled scope of a research study, by those who possess extreme, or morally unacceptable viewpoints. Recalling Gaver et al's *Energy Babble* device, (see section 2.4.4 for a detailed exploration of this work, [101]), that was created to produce a centralised, but varied, discourse around energy conservation. An extra source of content that could be drawn upon is those views that challenge the energy conservation narrative (the hegemony). Whilst not dismissing the existence of energy conservation as a cause for concern, different perspectives may challenge the means or motivation for energy conservation. An example of this is Lomborg's 2007 book *Cool It*, which challenged existing hegemonic perspective on climate change, and the unthinking adherence to reducing greenhouse gas emissions - as encapsulated by this quote: *"We need to remind ourselves that our ultimate goal is not to reduce greenhouse gases or global warming per se but to improve the quality of life and the environment"* [146]. In allowing such viewpoints to be presented, a device such as Energy Babble presents an agonistic interface because it allows for alternate, and conflicting discourse to be presented to the listener (this conflicting nature acknowledged by Gaver et al. in their paper).

I am arguing here that systems as such do not, and should not, attempt to take a neutral stance (which is a stance itself). As we can see from the previous example, by facilitating conflicting viewpoints, and by encoding that into the technical systems themselves, platforms are adopting an agonistic perspective - that debate should be *inclusive* of all viewpoints, for the good of democracy. As Thompson notes, everything around us is influenced by programmers and developers - they *"determine what products get created, what problems get solved, and what constitutes a "problem" in the first place* [...] *they're the architects. The decisions they make guide our behaviour"* [234, p.11], thus they encode their own views into the systems they create, along with their own assumptions of what is and isn't acceptable behaviour and/or viewpoints. This however often encodes the status quo, and in the case of many social media platforms and digital technology, the liberal ideology, which views sharing of views as positive, with the expectation of rational debate and consideration of ideas [9, 229]. Previous work (as discussed in Chapter 2) tells us that whilst this is the intention, it does not materialise in reality. The social factors invoked by these designs can also contribute to problematic discussion, for example the quantification and amplification of successful post, which uses *likes, shares* or *engagement* to identify content users are engaging with. This can encourage views that are more extreme, and thus evoke a stronger reaction from users, and in turn more interactions on a post (as explored by [36, 238]). Indeed people exhibit behaviours unique to online publics, as described by Suler with *the online disinhibition effect*, where people feel freed in their behaviour online. Whilst sometimes beneficial, this can also take toxic forms where *"people may be rude, critical, angry, hateful, and threatening, or they visit places of perversion, crime, and violence - territory they would never explore in the 'real' world"* [226, p.184]. It is therefore crucial for the designers of systems to acknowledge the perspective they are taking in their design, and what sort of debate will be facilitated by the design.

Here I draw upon the work of Slavoj Žižek, who discusses the positioning of liberal ideology against the idea of morality deeply rooted in culture. In order to reach consensus, those in society must put aside all of their morality and accept the liberal ideology - that is to say they should not have irrational moral objections (or antagonisms) that they cannot put aside. However, he argues that without morality, the boundaries for what is acceptable behaviour become endless:

> *"Without any 'organic' social substance grounding the standards of what Orwell approvingly referred to as 'common decency'* [...] *the minimalist program of laws which should just prevent individuals to encroach upon each other (to annoy or 'harass' each other) reverts into an explosion of legal and moral rules, into an endless process of legalization/moralization called 'the fight against all forms of discrimination.' If there are no shared mores that are allowed to influence the law, only the fact of 'harassing' other subjects, who - in the absence of such mores - will decide what counts as 'harassment'?"* [256, sec. II]

Thus, morality is still a critical part of society, as it guides people's reactions in contentious and important debates, and mobilises a strong reaction from people, which they can use to enter into debate. Rather than suppressing contentious debate, agonistic interfaces thus frame and facilitate it in spaces where it can be conducted in a respectful, and ultimately useful, way. Indeed, striving to create a society where there is no conflict (as Žižek explains) would reduce debate, as it *"would be on this very account a society condemned to see crimes everywhere."* [256, sec. II]. Therefore we need morality, and moral reactions, to motivate people to engage in debate and discussion, to refute points with which they do not agree, as part of an agonistic struggle.

7.7 Limitations

It is important to acknowledge the limitations of this work, and I address the specific limitations of the individual studies presented in Chapters 4, 5 and 6, and also exercise reflexivity on the approach of the thesis more broadly.

7.7.1 Moral Compass & Spotting Guide - Workshop Format

Both the Moral Compass study (Chapter 4) and Spotting Guide study (Section 5.2) were conducted using a similar format, consisting of three workshops spread across three weeks, with a homework task after workshops one and two. Using this format created a controlled setting, which whilst useful for data collection, may have had an

impact on participants' responses. In particular, the reflection and criticality offered may be heightened by participants thinking forward to the face-to-face workshop sessions more so than in normal day-to-day usage. Cognisant of this I iterated on my methodology for data collection and the overall deployment of the prototypes, and allowed the participants to use the prototype in their normal day to day usage, as discussed in Chapter 3

7.7.2 Screenr - Scaling of Co-Criticality

The prototypes I developed have tended towards deployment with 10 to 15 participants, either in controlled circumstances through weekly workshops, or deployed over the period of one month. It is important to acknowledge that they have not been deployed with large groups of users, and that the findings reflect that. However, I have framed my discussion accordingly, and where possible suggest how any of the prototypes might be integrated into a larger system, or developed and deployed with a large numbers of users. Throughout the review process for the publications comprising this work, I have helpfully been encouraged to consider and address how these prototype systems might work when deployed in the real world.

The Screenr study (Chapter 5) presents some limitations that must be discussed, particularly how a system relying on co-produced criticality would work with more users. I purposefully used a small number of participants, not only because it was a technological prototype, but at larger scale, the quality of interaction changes, and in the context of critical co-viewing, I would hazard that it degrades somewhat. Specifically, an increased number of users in a real time chat would inevitably lead to difficulty keeping track of conversation given the increased number of messages. The Screenr study demonstrates the group size was small enough to allow for 'closer readings' of the TV content in a small, safe space. The delicate balance between creating tags, and then using chat to contextualise and explain their views would also be affected in larger groups. Likewise, the social cohesion and sense of social accountability demonstrated through the co-viewing process would reduce, leading to less meaningful voting choices and a more random viewing schedule. Furthermore, I also restricted programme choice during the study for logistical reasons. There are a finite number of reality TV programmes in a TV schedule each week, and in

order to maintain a good number of concurrent live viewers each week I reduced the selection, to allow for more decisive voting. In the future work section (Section 8.2) I discuss how these limitations can be overcome to produce agonistic interfaces that engage a large number of users and account for real world variation and scheduling pressures.

7.7.3 Criticality of Reality TV

Given the socio-political focus of reality TV, the form of critical reflection and type of reality TV watched will differ both within a country (region by region) and across countries. For example, the poverty porn studied in Chapter 5 is a UK-centric sub-genre of reality TV, and therefore focuses on socio-political topics that have meaning to people from the UK, such as people claiming state welfare, or minority populations such as the Romani or Traveller communitites [74]. It is also common for reality TV producers to adapt their programme for individual cultures/countries, thus tailoring the socio-political issues to fit those of the culture (see Bignell's discussion of the worldwide *House* reality TV series [23, p.80-81]).

7.7.4 Spkr - Coding of political classification

In order to operationalise the Nolan chart political alignment used to align participants and content in the Spkr study (Chapter 6), a substantial amount of labour was required, most notably for identifying relevant topics on a daily basis and coding the tweets associated with that topic. The coder also needed to be familiar with Meek's [158] political alignment questionnaire in order to classify the content. For my study I did not have access to other researchers for assistance, but a way that future work could reduce the labour required would be having multiple researchers. Not only would this reduce the individual labour, but it could be used to generate inter-rater reliability measures. As I discuss in Chapter 6, whilst not only being time consuming, this process involves some subjectivity, albeit somewhat mitigated by use of classification from the Nolan chart and Meek's questionnaire. Generating inter-rater reliability, made possible by having two or more coders, would be a further way to control for this subjectivity, if desired. There is also an opportunity for the use of AI

powered automated classification, that could be trained by human coders. Also, it should be acknowledged that by using Meek's version of the Nolan chart, the work is UK-centric. It is possible to use a different version of the Nolan chart, specific to the political context being studied, such as the US-centric version originally presented by Nolan [183]

7.7.5 The Agonistic Perspective of Democracy

Moving onto more broad limitations of the thesis as a whole, it is essential to acknowledge that throughout, I take a specific view towards ideological diversity as part of democractic debate, which is agonism. I have discussed (in Chapters 1, 2 and 3) why I take this stance, and the background of agonism in design work. However it would be entirely possible (and indeed scope enough for another thesis in itself) to take a different conception of democractic debate and to design with that in mind. To illustrate, Bozdag and van den Hoven explore how the problem of the filter bubble and echo chamber can be viewed in different ways depending on the conception of democracy being used [34]. As noted in Section 2.3.1 on page 29, to the agonistic perspective a filter bubble removes opportunities for conflict. By contrast, to the *contestatory* perspective of democracy, which maintains citizens should be allowed to contest any decisions not made transparently, *'filter bubbles prevent awareness of both the items that people could disagree with and the information on the basis of which they could justify their reasons for disagreeing"* [34, p.253]. Whilst this is a limitation, it is also presents fertile ground for future work that adopts a different conception of democracy and uses it to address the problems of online publics and technology design's role in facilitating these problems.

7.7.6 Limitations of studying 'the public'

This thesis is framed around problems with online publics, and has presented studies around three domains of inquiry that aim to design digital prototypes that overcome or reframe these problems. The number of participants in my studies is relatively small, in line with my critical technical practice methodology, which allowed me to examine in rich detail how participants interacted with the digital prototypes. Of

course, this is also constrained by funding pressures - to compensate participants for their time, to pay for equipment and services - as well as being restrained, purposefully, to remain within the scope of a PhD thesis. As such this work is limited in that it does not interact with a wider public, and therefore I have not worked with the huge diversity of viewpoints and opinions that are held across society. Without engaging with this full range of viewpoints and opinions it is difficult to know how the digital prototype would have been used, and what form the resulting discussion, reflection and criticality would have taken.

However, given the constraints of working at the scale of 'the public', the work in this thesis still provides insights on how users, at an individual, and small group scale experience agonistic interfaces, and how that effects the way they engage with socio-political discussion. In my discussion I have drawn out how their usage differs from existing practices, and in what ways this changes the nature of socio-political discussion. As demonstrated in the literature review, other work in this area has engaged with similarly sized groups of participants (e.g. Gaver et al.'s Energy Babble [101], and Gorkovenko et al.'s Social Printers [107]). There are examples of HCI work that has studied the public at larger scale, but this approach provides sparse detail about participant's experience (e.g. Kriplean et al.'s ConsiderIt engaged 6000 [137], and Facebook's study of emotion in the news feed algorithm, using over 600,000 participants [136]). The difficulty of studying publics in rich detail has been highlighted by Gaver et al., who deployed 140 of their Datacatcher objects. They note that *"large scale trials of computational products have rarely reported [sic] in the CHI literature, and were of limited assistance in planning a study for diverse and disparate participants"* [100, p.1599].

7.7.7 Deploying Prototypes In Situ

How participants are disposed towards research devices, and how they act around them has been the focus of recent attention by Tolmie [235] who explores the range of ways that research teams deploy their research prototypes, paying particular attention to how the researchers are positioned towards how the prototype will fit into the participants environment. Even the presence of a device for research purposes

changes the participants relationship with the device: *"once people orient to a deployment as being about 'doing research', this provides a to-hand account for setting aside any of the usual assumptions"* [235, p.224]. Of relevance to this work, Tolmie notes the way participants will privilege the research device more than they would a device they bought and would install themselves, such as where they position the device. Furthermore, the ownership of the device (by the research team) sometimes creates interesting dispositions to the device, for example relinquishing all maintenance of the object, or being overly protective of the device. Whilst it is difficult to assess to what extend these factors were manifest during the deployment of Spkr in participants' homes, it is not unreasonable to expect Spkr was granted some privileges in its placement in the household, and that it was treated as a research prototype, not another piece of household technology. However, it is clear from the results that despite it's somewhat special status, for some participants it did become part of their household rhythm and became something they expected to listen to during the day.

7.8 Summary

My inquiry was guided by a set of research objectives, and in this chapter I described how and where in this thesis I have responded to those objectives. Furthermore I demonstrated how the studies respond to the three research questions outlined at the beginning of this thesis. I described how interfaces can invoke agonism, as evidenced in my literature review and methodological underpinnings in agonism, adversarial design and critical technical practice. I then synthesised my results to discuss how agonistic interfaces can be designed to create and promote critical reflection and engagement, and I furthered this discussion to examine what forms of discussion, reflection and criticality were evoked by the agonistic interfaces. In the following chapter I discuss the conclusions I have drawn from this work, and then I present my ideas for future work, comprised of a broad research agenda, where agonistic interfaces may *not* be appropriate, and finally some future agonistic interface prototypes that could be built.

Chapter 8

Conclusions & Future Work

“In what ways can interfaces invoke agonism?" This was the first of my research questions, which I responded to by examining the ideological underpinnings of digital platforms, and discussing prior research that demonstrates how different conceptions of democracy, such as agonism, can address some of the acknowledged problems with online publics and digital platforms. I have used the term agonistic interfaces to talk about these digital systems that do the work of agonism - the specific sites within digital technology where users can be engaged in agonistic debate and reflection around socio-political issues.

Based on this foundation, the second guiding question was *“How can agonistic interfaces be designed to create and promote critical reflection and engagement with socio-political topics?"* Through four studies I designed agonistic interfaces, based on an analysis of the existing technology design practices around the three interesting areas: i) socio-political social media, ii) second-screening of reality TV, and iii) conversational interfaces. These studies generated tacit knowledge of the specific design features that can facilitate and sustain agonism. These range from pushing content onto users and adding friction to their user experience, through to invoking social responsibility and capitalising on live interaction.

Beyond the design elements of agonistic interfaces, my third question asked: *“What forms of discussion, reflection and criticality are evoked by agonistic interfaces?"* I have shown that agonistic interfaces can work to concentrate various discourses in the same place, enabling reflection of a broad spectrum of content not normally considered together. We have seen how agonistic interfaces can pro-

vide the stimulus to initiate discussions within a family environment, or with others through the agonistic interface itself, and that processes of self-censure and a user's internal bias become a factor when exercising criticality.

In order to carry out the research work required to respond to these questions, I adopted a critical technical practice (CTP) methodology, which guided me to critically examine the existing assumptions and practices of technology design. In doing so, CTP helped me to re-design technology based on different assumptions, or to challenge existing practices. Combined with CTP, I used adversarial design - a methodology grounded in agonism - to build a suite of digital prototypes that invoked agonism through engagement with socio-political issues.

Throughout I have discussed three interwoven threads: designing agonistically-informed interfaces, designing to support democracy, and designing to support a plurality of viewpoints. These threads are all related under the umbrella of making agonistic interfaces - that enable agonistic debate, and thus support democracy by addressing problems of viewpoint diversity and engagement with debate in online publics.

Over the course of the four years studying for this thesis I have engaged with 50 participants, who have in turn contributed approximately 1,423 days in total, across all of the studies. Analysis of their valued input and use of the digital prototypes has led to the publication of three papers at the international peer-reviewed ACM Human Factors in Computing Systems (CHI) conference. The Spotting Guide (Section 5.2) and Moral Compass (Chapter 4) studies were presented as a paper at the CHI 2017 conference in San Jose, CA, USA. The Screenr study (Section 5.3) was presented at the CHI 2019 conference, held in Glasgow, UK, and the Spkr study (Chapter 6) has been accepted for publication in the CHI 2020 conference, and will be presented in Hawaii, HI, USA in May 2020.

8.1 Contributions

This thesis tells us how we (the research and practitioner community) can specifically design digital systems to enrich and diversify the debate, reflection and criticality that users engage in. It demonstrates how to identify and build interfaces, that I

term agonistic interfaces, which reconfigure or change the way people engage with socio-political issues.

Domain one: Socio-Political Social Media This chapter demonstrates how to reconfigure the mode of consumption for a social media stream, and how this can be designed to prompt reflection on the contents of the stream. The Moral Compass study demonstrates how morality can be used as a lens to view socio-political social media content.

Domain two: Second-Screening of Reality TV This chapter presents second-screening activities as a means to prompt reflection on problematic TV programmes, using *lean forward* activities that engage users with the programme, as well as facilitating socially constructed criticality of the programmes. It also documents how agonistic interfaces can be designed to facilitate group criticality and reflection on live TV programmes.

Domain three: Conversational Interfaces This chapter demonstrates that smart home technology can be used as a means to inject socio-political discussion topics into the everyday context, and precipitate debate and discussion within the family environment. It also contributes an implementation of the Nolan chart of political alignment as a means to introduce viewpoint diversity, increasing exposure to different viewpoints on topical socio-political issues.

Overall, I present two primary contributions in this thesis:

1. identify agonism as a means of addressing existing problems around engagement, exposure and debate around socio-political topics, and describe agonistic interfaces as an approach to digital technology design to address this.

2. demonstrate how agonistic interfaces can function in specific contexts to reconfigure debate and engagement with socio-political topics, through the design and deployment of four technology prototypes. Grounded in the existing design practices in each domain, as well as my own critical technical practice,

these studies elaborate how to design and deploy agonistic interfaces, as well as contributing new knowledge to the domains of socio-political social media streams, second-screening of reality TV, and conversational interfaces.

These two contributions thus extend knowledge of the research gaps and opportunities around diversifying and enriching socio-political debate and criticality in online publics, and how to tackle these concerns through the design of digital technology.

8.2 Future Work

Casting an eye to the future, what research directions have been opened up by the contributions of this work, and what are the gaps in knowledge that could be explored? As you will have seen over these concluding pages, there are some limitations to the studies I have conducted, but some of these highlight immediate opportunities for future work. Moreover, studying agonistic interfaces in situ has demonstrated where agonistic interfaces are useful, as well as where they might not be. As a starting point, and to frame the remainder of this section, I would like to consider where agonistic interfaces may *not* be appropriate.

8.2.1 Where Agonistic Interfaces May Not Be Appropriate

It would be remiss of me to include discussion of future work without discussing areas that I think based on the findings and discussion, would *not* benefit from an agonistic interface, in as much as there are ethical or moral tensions (that I have highlighted throughout this thesis, and will signpost again here), that would make their implementation difficult, or problematic. As a caveat, I am not saying that one should never conceive of creating an agonistic interface in such a context, but more that careful consideration should be given, as these contexts do not seem to naturally lend themselves to agonistic debate, as I will explain:

Non-contentious issues Agonism operates around the assumption of conflict, and that people engaged in discussion will eventually reach a point where they will not agree. Therefore, as discussed earlier in this chapter (Section 7.6.5), introducing agonistic debate into topics that are not contentious - that are decided by society and *morally* do not wish to be debated - would most likely cause users to disengage. For example, murder is framed as morally bad, and laws enshrine this moral stance. You could challenge this stance, but people may chose not to engage in such a debate.

Interfaces for utility As I have shown, interfaces can be used for a variety of purposes, but it is important to remember some interfaces are used for a purely functional reason, and as such come with a different set of priorities placed upon them by the user, and the developer. Social media interfaces focus on the social factors that bring people together, encouraging them to share and consume content. An online banking interface by contrast focuses on the privacy and security factors, which the user expects to be paramount. An agonistic interface for online banking may be successful in challenging the hegemony (*"is your pension helping the bank give loans to problem gamblers?"*, *"can one ever truly own money?"*) but is unlikely to be the favoured means of quickly administering a bank account. Again I am not saying we should not think about such a thing, but rather there are considerations for those using an agonistic banking interface, and also for the company which decides to introduce it.

Non-democratic contexts Throughout I have discussed how agonism is founded on the idea of democratic struggle, and specifically conceptualises that this should be accommodated by inclusive, conflict-filled, debate. However, not all societies / cultures are founded on the idea of democracy, and thus this limits the ability for agonism to provide inclusive debate, given non-democratic ideas such as political censorship (e.g. the censorship of the Internet by the Chinese government [143]).

8.2.2 The Future of Agonistic Interfaces

Having considered some tricky or complex contexts, the remaining question is: what are the interesting and feasible research directions presented by agonistic inter-

faces? Over the course of this PhD a number of significant societal events have occurred, and phenomena that were nascent at the outset of this PhD are now acknowledged problems with online society, such as the polarisation of online publics. The Brexit Referendum in the UK, followed shortly thereafter by the election of Donald Trump as US President are still being discussed in terms of the echo chambers that each side resided in (e.g. [119]). Added to this the continued rise of populism in Europe and terrorist attacks founded on hatred of the Other (e.g. see Plambech for a detailed description of Othering around the Christchurch terrorist attack in 2019 [192]), it is clear that the motivation for this thesis is as relevant now as it was at the outset.

A future research agenda would look to continue examining how we, as technology designers, can bring people together for discussion, to share, rebut and debate their views. At the very least, as I have demonstrated, this involves putting aside existing models of democratic debate (liberalism and libertarianism) and embracing agonism - embracing the reality that we all have diverse views, and it is likely on some issues we will never agree. I suggest some broad directions that may be useful here, and in the following section explore them in more detail.

Leverage new technologies In the Spkr study (Chapter 6) I studied a new technology, conversational interfaces within smart home devices, and used the powerful mechanisms afforded by them to introduce reflection and engagement with socio-political issues. Applying a similar approach - examining the powerful mechanisms with other new technologies and the assumptions in their design process - could prove fruitful. For example, virtual reality (VR) technology is currently used primarily for gaming, as well as training and experiences, but it could be used to engage users in face to face debate.

Identifying contexts Focusing on applying agonistic interfaces to new contexts where socio-political topics are relevant is a useful direction. To illustrate, the Spotting Guide and Screenr apps were relatively simple smartphone apps, but their use in the context of problematic TV programmes, as part of the TV viewing experience made them useful at engaging criticality and debate. Another space that concerns

the socio-political could be public transport. Newspapers are often found here, and historically buses and trains were places where strangers were apt to pass the time with conversation, prior to the ubiquity of smartphones.

Supporting agency The agonistic interfaces I've discussed have allowed users to exercise criticality and reflection in their own way, by supporting them and giving them agency, whilst being within a structure that encourages them to be reflective or critical. For example the Moral Compass allowed users to split the Twitter stream based on their own subjective views, and this could be applied to other contexts such as online news, exposing news bias and allowing users to manipulate and tweak the biases of their own recommender algorithm, something that is advocated by Caplan and boyd [49].

8.2.3 Future Prototypes

To move beyond abstract discussion of research directions, I have drawn on these research themes to present six interesting opportunities for future work in the form of digital prototypes, similar to those I have presented throughout this thesis. Not only is this to articulate the future directions more concretely, but also to act as direct actionable inspiration to other researchers who may wish to take this work forward. I have also suggested the research question (in *italics*) that each future digital prototypes might answer.

Critical Viewing Collectives

How might small-scale critical viewing translate to the large scale platforms such as BBC iPlayer and Netflix? It is clear that bringing TV viewers together using a digital prototype that encourages and supports their own critical viewing processes, in their own terms, has been positive in allowing the viewers to share ideas, and to express things in a safe space. I see the maintenance of the 'small safe space' idea imperative to the success of the critical co-viewing process as described in Chapter 5. Taking this forward, future work could look at ways to create these critical viewing collectives as part of a larger system.

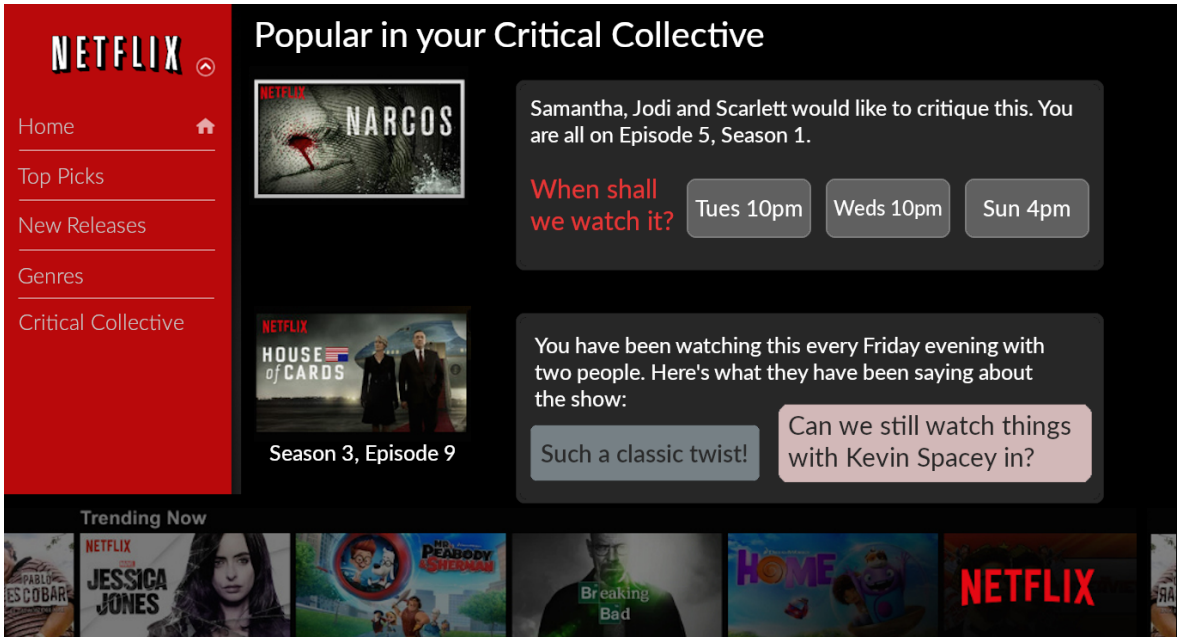

Figure 8.1: Mock up of The Critical Viewing Collective interface, as part of the Netflix interface.

Taking Netflix as an example, given the extensive range of data they collect about users, they know who is watching what at which times, and have an understanding of the patterns of viewing, including sub-patterns relating to the genres of programmes a viewer watches at different times/days. Using this information, Netflix could create a *Critical Viewing Collective* by grouping users who regularly watch together (*together alone* - as they do not know others are watching at the same time). In the spirit of agonism, it would be imperative to find diverse viewing patterns within those who regularly view together. For example, bringing viewers who regularly watch reality TV and documentaries about the economy, together with viewers who regularly watch reality TV and soap operas. Given the abundance of user data on platforms like Netflix, these algorithms could be refined further to bring together viewers with diverse, but overlapping, viewing patterns. Figure 8.1 shows a mock-up of what the Critical Viewing Collective might look like on the Netflix interface. Users would be invited to enter the chat, after being given a welcome introduction that explains they have been watching TV with these people for the last 6 months without realising it, and that Netflix think they should get to know them.

In the same way as bringing a group of unrelated participants together in a digital prototype (Screenr), the Netflix Critical Viewing Collectives could invite users to participate in a critical exercise, at their own volition. Whilst it may not be for every user, as I found during the Screenr study, it also offers inquisitive users an opportunity to experience critical co-viewing, and in the same way that Screenr participants found, may leave an impression on them as viewers beyond using the app.

Importantly, the capability to build such a system already exists, as can be seen if looking at the Facebook Live feature[1], which allows you to co-view (although without any support for criticality) Facebook videos together. Similarly, YouTube provides a 'Premiering' feature[2] for content creators, meaning their videos are viewed simultaneously by all viewers, putting them all in the same chat window. To reinforce my rationale for keeping groups small, an example can be seen by looking at Twitch.tv channels, which place thousands of viewers in the same chat channel, which quickly becomes overwhelmed with the number of messages, and becomes largely incoherent (as evidenced in [180]).

Social Media Moral Observatory

Can we produce large scale alternate interpretations of social media data streams, such as by moral reactions? Moral Compass produced moral coding of Twitter, and as we have seen was useful for framing debate and reflection about the content of the Twitter stream. Moral Compass relied on manual coding by participants, which is valuable as it was able to draw on individual subjectivity. The *Social Media Moral Observatory* (SMMO) scales this up. Users volunteer their own socio-political perspective through an assessment test and then provide some training data by coding a set of tweets by morality. Then the system underpinning the SMMO feeds this training data and assessment data to the machine learning algorithms which then automatically classifies Twitter data. The SMMO therefore creates a different dimension on trending topics. In the existing model used by Twitter topics can be trending because they are popular, or equally they can trend because the content is unpopular and lots of people are talking about it. Trending topics are also liable to manip-

[1] `https://www.facebook.com/facebookmedia/solutions/facebook-live`
[2] `https://support.google.com/youtube/answer/9080341?hl=en`

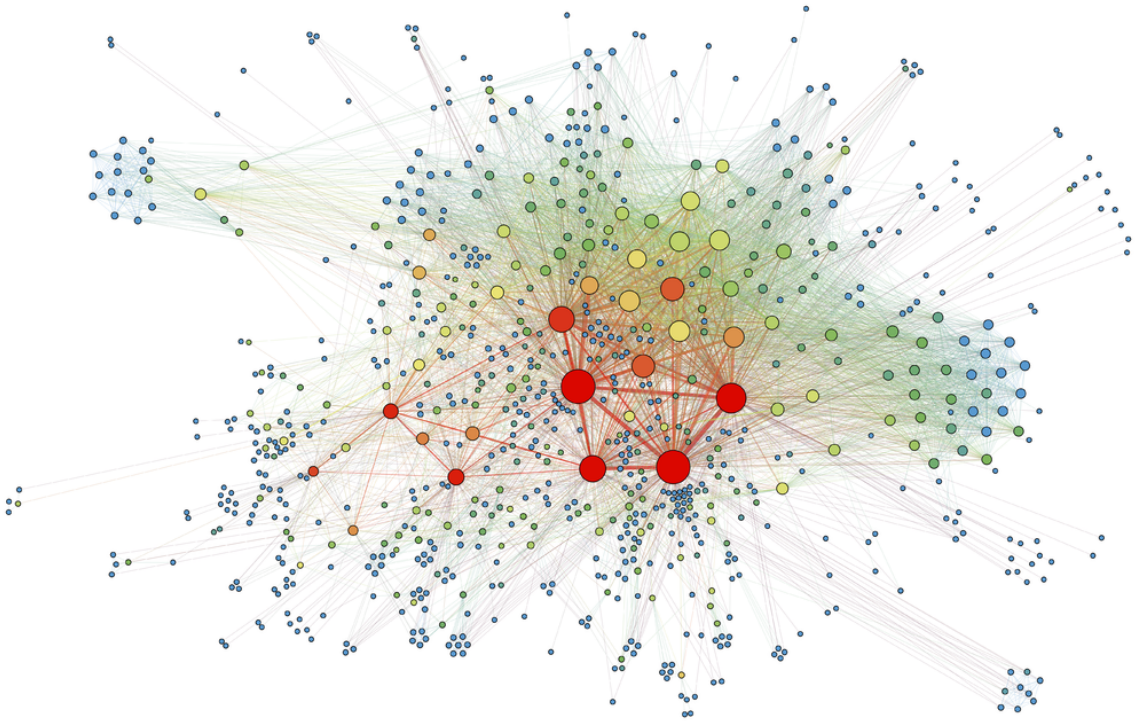

Figure 8.2: Example of how The Social Media Moral Observatory might look, showing a coding of moral reactions to the live Twitter feed. Based on: *"Social network visualization" by Martin Grandjean. CC-BY-SA 4.0.*

ulation [87]. Without reading the hashtag stream, it's not possible to know what the reaction to a trending hashtag is. The SMMO would be a way to explore the social media reaction, and given the flexibility afforded by automation, the SMMO allows users to view multiple different moral codings of Twitter, made by multiple individual users.

The SMMO is measuring an important metric that is currently under-monitored. For example, by considering the moral reactions to tweets, it's possible to see whether a tweet has gone viral. Whilst this may seem trivial, when taking the moral reaction component into consideration, it can be used to identify nascent discussions that may lead to abuse. In effect, the SMMO is able to detect public shaming, as a user of the SMMO would be able to see an overwhelming negative reaction, such as *scorn*, or *outrage* around a particular tweet. Public shaming via social media is a growing problem, and as discussed by Ronson [210], a practice that has been

relegated to history, because society called for compassion for victims of unfettered public shaming. Yet this is largely unchallenged on social media. For example, the SMMO would have been able to identify the overwhelming negative reaction rising around the tweet of Justine Sacco in 2014, who became the number one trending topic on Twitter after a joke was taken out of context whilst she was on an 11 hour flight, and she was publicly shamed across the world, with many negative effects on her life [22, 210]

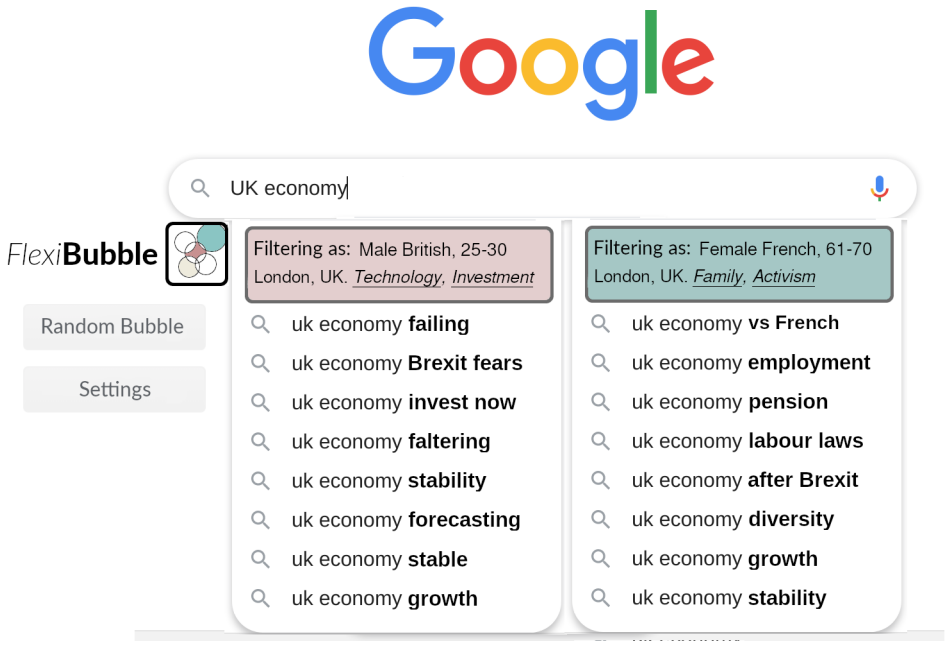

Figure 8.3: FlexiBubble showing two filter bubbles from the Google search bar.

Flexi-Bubble

How do you know if you are subject to a filter bubble? What would your feed look like if it was showing a different echo chamber? Participants during the Screenr study noted that applying critique to news coverage would be an interesting activity. *Flexi-Bubble* is an extension for most web browsers that uses an advanced social-media scraping technique to allow the user to interactively apply echo chambers to their Twitter and Facebook feeds. Whilst browsing social media, Flexi-Bubble can

be toggled, and depending on the topic you are viewing, will give you a dial which can be turned, with the result that the content you see in your feed is drawn from the echo chambers you specified. In order to function correctly, Flexi-Bubble needs access to your account, so that it can locate you within the network and identify your existing position within the echo chamber(s). Flexi-Bubble also works with search engines, such as Google, TripAdvisor and AirBnB to allow the user to interactively change the filter bubble they are currently subjected to (Figure 8.3). As a result the user is able to see whether a website is using their data - their current location, the time of day, device - to filter their data. The interface also allows the user to toggle various combinations of filter on, to see the effects.

The Curious Smart Home Device

When designing and implementing Spkr, I intended the device to remain somewhat neutral, as it would speak out the diversity of viewpoints, rather than discussing it's 'own' views. As was highlighted by one participant, given that the device had a distinct voice, and the form of the Twitter content meant it often used the first person, to some participants Spkr took on a personality of it's own. *What would a purposefully provocative, opinionated smart home assistant look like?*

A further iteration of Spkr, *The Curious Smart Home Device* is intended to provoke those around it by making statements that go against the views of those in the household. Rather than being a smart home *assistant*, the Curious Smart Home Device requires educating by those around it in order for it to understand how the world works. Using machine learning and natural language processing, The Curious Smart Home Device is able to understand the responses from those in the household, to encourage users to explain, in their own terms, the vagaries of their particular political viewpoint. For example, the Curious Smart Home Device might ask *"why is migrant labour bad for the NHS?"*, which would illicit a response from those around it, based on their political views (e.g. *"it isn't bad for the NHS, but it can be bad for other parts of society"* or *"it indicates there's a shortage of UK educated NHS doctors"*). The Curious Smart Home Device is connected to the news sources and social media, which are processed by the machine learning algorithms, and anything it does not understand will be formed into a question. In form it is

a playful device, that encourages those in the household to help it understand the world, and by questioning seemingly obvious points, or requesting complex concepts be explained to it, it attempts to prompt reflection among those in the household. As noted by Sagan: *"There are naive questions, tedious questions, ill-phrased questions, questions put after inadequate self-criticism. But every question is a cry to understand the world. There is no such thing as a dumb question"* [212].

Figure 8.4: Players sitting in the study in vTime XR - Café de Paris might look like this when players are engaging in heated debate. *Copyright 2020 vTime XR, vTime Limited.*

Café de Paris

How might we reduce disinhibition and encourage supportive debate among people? Social VR applications are designed to get people to interact and to talk with each other, as well as share experiences. For example social VR games such as vTime[3] allow groups of players to share stunning sunsets together, sit around a campfire toasting marshmallows or travel to the north pole together. In a recent

[3]vTime XR. `https://vtime.net/`

study McVeigh et al. interviewed developers and content creators for social VR games, who discussed how connotations of the environments affected the way users interacted (one developer avoided the locker room environment, given its association with crude “locker room talk"). The developers also describe the use of props, objects and activities that can be introduced into an environment to act as a *“social lubricants"* [157, p.546]. *Café de Paris* is a social VR application to bring groups of players together in a relaxed café setting, with the intention of users engaging in conversation, people watching, or simply quietly relaxing. Drawing on the cultural conventions of the French café as a space for debate, this social VR application encourages debate and conversation between users, face-to-face, with props such as newspapers, that players can read through, which contain automatically generated topics for debate (such as contentious news topics). There are also computer controlled characters, such as the bartender, who initiates debate and reacts to the actions of the players in the café, as well as to what they are saying. A planned expansion for Café de Paris will allow players to experience Othering as part of the game, where characters will react negatively to the player character, and some services will be unavailable or difficult to access.

The Agonistic Broadsheet

Can traditional offline formats be used for agonistic interfaces? Newspapers are a historically important medium for the consumption of news, with their pivotal place at the centre of the coffee houses of Western Europe following the widespread uptake of the printing press in the 17th century. They are still a dominant news source in many countries across the world, despite the increasing shift to online journalism. The physicality of a newspaper also enables other interactions with them beyond simply consuming news. For example, those in the household may perform graffiti or annotation on the newspaper, thus impacting the way news is consumed by others in the house [251]. Furthermore, newspaper has some societal uses, for example in the UK being used to light barbecues and fireplaces, to wrap items or line pet bedding, and it can be accidentally read during the process, piquing interest in old, possibly outdated, long forgotten news topics - something I have experienced personally. These qualities are embodied through the *Agonistic Broadsheet*, which

Figure 8.5: The Agonistic Broadsheet smart home letterbox, in the process of printing the days edition of the Agonistic Broadsheet

is a personalised daily newspaper printed by a smart home letterbox fitted in your front door (see Figure 8.5). Based on your online reading history, the Agonistic Broadsheet is printed each morning by your letter box, and the content is algorithmically populated with diverse opinions of the news topics that you have been reading. The smart home letterbox also incorporates OCR technology, allowing you to write a letter, or annotate the agonistic broadsheet, and put it back through the smart home letterbox. It will scan the response and, following a similar process to the traditional *letters to the editor* means of responding to newspaper articles, your letter may be published in other people's copies of the broadsheet, allowing you to engage in agonistic debate via a printed newspaper.

8.3 In Closing

As technologists - designers, researchers, engineers - it is clear that the technology we conceptualise and put into use has effects: on users, communities and society. I have presented agonistic interfaces as an approach for the design of digital systems that value inclusivity of viewpoints, and maintain the idea that views *will* conflict, if we are to engage in any sort of emotive democratic debate. I hope that this thesis, and the agenda for future work that prompts criticality, debate and reflection on important socio-political topics proves inspirational to others in the research community, and that together we can continue to bring users face-to-face with surprising, challenging, exciting and provocative ideas and viewpoints that will ultimately enrich society.

I would like to conclude with this quote from Daryl Davis, life-long anti-racism activist, who sums up why debate, even when filled with conflict, is imperative for the good of society:

> *"If you have an adversary, you don't have to respect what they are saying, but respect their right to say it, and have that conversation. We spend too much time talking about each other, at each other, past each other, and not enough time talking* with *each other* [...] *When two enemies are talking, they're not fighting, they're talking. They might be yelling and screaming, but at least they're talking. It's when the talking ceases that the ground becomes fertile for violence, so keep the conversation going"* [62].

It is my hope that agonistic interfaces will continue to encourage respectful conflict between parties, and in this way work to keep the conversation going.

Appendix A

The Moral Compass Participant Information Sheet

Invitation to take part

Thank-you for expressing an interest in this research project. You are invited to take part in a research study. Before you decide if you will take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and to discuss it with others if you wish. If there is anything that is not clear or if you would like more information please contact the researcher [Email address here]. Please read the below consider whether or not you would like to be a participant - if you are happy to be involved please reply to the email address provided with your copy of the questionnaire. Thank you for your time and consideration.

What is the purpose of the research?

Our project, CuRAtOR (www.curator.ac.uk), is exploring how digital applications might be developed to positively counteract Othering (the practice of making distinctions between groups of people, often negatively) in online environments. We have developed a digital “companion" application, The Moral Compass, to be used as a way to organise and promote critical readings of information on Twitter. We are interested in finding ways to develop digital technologies that can help people identify, and counter online Othering, encouraging online environments where people

are free from stigmatisation and marginalisation in daily life.

In this research, our aim is to identify whether our application can help identify and counter the practice of Othering and stigmatisation that may occur on Twitter and on social media generally. We would like you to use our application to view a selection of tweets and attend three 60-90 minute workshops at the University of Bath to talk about your understanding of online Othering and how you have used our application. This will enable us to gain insight into what aspects of our applications design people have found useful, enjoyable, and insightful, along with identifying areas for it to be improved as a tool for exposing and countering stigmatisation and marginalisation that occurs online.

Who is doing this research?

This research will be conducted by various members of the CuRAtOR research team, and supervised by the CuRAtOR principal investigator, Prof. Shaun Lawson. All members of the research team will work to ensure that the research abides by professional ethical standards.

Do I have to take part?

No. Participation in this trial is on a strictly voluntary basis - you do not have to take part in the trial and can withdraw from the trial at any point, without providing a reason, should you choose to do so.

What will I be asked to do?

We would like you to trial a web application developed by us, called The Moral Compass, which will involve attending three 60 minute workshops at the University of Bath campus.

In the first workshop, you will be invited along with 3-5 other participants, to hold a discussion about your existing understanding of online Othering. We will then introduce the Moral Compass application and run through some examples of how to use it. You will then be asked to use the application to label (i.e. code) and categorise a selection of tweets pre-defined by the researchers, then reflect on the

process with the wider group. You will be provided with a £10 voucher (of your choice) at the end of this workshop as a token of appreciation for taking part in the study. You will also be asked to use the application in your leisure before the second workshop, to code tweets related to two designated cultural events (i.e. television shows) taking place that week.

At the second workshop, you will use the application to look at a selection of tweets coded by the group at their leisure before the session, and reflect on the process with the wider group to explore discuss differences in codes and interpretations. At the end of this workshop you will be asked to do two activities: 1) code a pre-selected Twitter hashtag around a TV show being broadcast live that week, and 2) engage with one of your fellow participants Twitter codings in relation to this event.

At the final workshop, you will discuss your tweets and codes and explore a visualisation of yours and others' "Moral Compass" (a visualisation of how you and your fellow participants have coded and interpreted your tweets). At the end of this workshop participants will be asked to reflect, as a group, on your experiences with the application. At the end of the final workshop, you will be provided a £30 voucher (of your choice) as a token of appreciation for taking part in the study.

Are there any disadvantages to taking part?

The study outlined here has been designed such that we do not anticipate any risks involved that would go beyond those experienced in everyday life. However, tweets that may be coded and read within the Moral Compass application may at times feature socio-political content which some participants may disagree with. In the remote eventuality that participants become distressed upon encountering this content, we will provide information about persons or groups who may be contacted that can provide support. Furthermore, we remind participants that they are free to withdraw from the study at any time and without providing a reason.

What are the benefits of taking part?

There are no direct benefits for those individuals taking part in the study. However, you will be taking part in world-leading research that is exploring ways that software

technologies can be used to make people think more critically and reflectively about social and political messages encapsulated in social media conversations. You will be providing us with vital insight about the ways such technologies could be used and designed in the future.

What information will you collect about me?

We will audio and video record each workshop session in entirety. This is required so we may review discussions that occur in the workshops. The recordings we make will not be heard or seen by anyone outside of the research team, and will be stored securely on a password-protected computer in the Computer Science department at the University of Northumbria, which is the academic base of the CuRAtOR project. We will also collect usage information from the Moral Compass application, so we can understand what and how you have used it. This information will include details of the tweets you code and the categories you create. Again, this information will be held on a password-protected computer in the Computer Science department at the University of Northumbria.

What will happen to information you collect about me?

Your privacy will be protected at all times. Your identity will not be known by anyone other than the people directly involved in the study. None of your personal details will be stored alongside the video or audio recordings and application usage information, and you will be given a code (for example P1, P2, P3) so that your identity will be protected. Any recordings of you will be stored securely at the University of Northumbria and will not be used for any other reason apart from the study. Information collected from using the Moral Compass application will be anonymous, and you will not be identifiable. This anonymous information will be used in publications about the study, such as academic journals and conferences. If you would like to be kept informed about these publications please tell the researchers.

What will happen if I do not want to continue with the research?

Participation in the research is entirely voluntary. If you decide you will take part you will be asked to sign a consent form. You will be given a copy of the information sheet and consent form to keep for your records. If you do not wish to carry on with the research you can withdraw at any time, without giving reason. Furthermore, you are entitled to request to withdraw any data you do provide for the study, for up to 30 days after completion of the study. After this point, your data will have been anonymised and as such, it will no longer be possible to identify you and remove your data.

Who do I contact if I have a complaint?

If you have a complaint you can contact the Principal Investigator for the study, Professor Shaun Lawson, using the details provided below:

Prof. Shaun Lawson Dept. of Computer Science and Digital Technologies, Pandon Building, 2nd Floor, 238, University of Northumbria, Newcastle-upon-Tyne, NE1 8ST Telephone: 0191 2273944 E-mail: shaun.lawson@northumbria.ac.uk

If you have a concern or complaint and wish to speak to a person who is independent from the study, you can contact the Head of Computer Science at the University of Northumbria:

Prof John Woodward, Associate Dean (Research and Innovation), Faculty of Engineering & Environment, University of Northumbria, Newcastle, NE1 8ST Telephone: 0191 227 3048 E-mail: john.woodward@northumbria.ac.uk

Who is organising and funding the research?

This work is funded by: Economic and Social Research Council (ESRC); Arts & Humanities Research Council (AHRC); Engineering and Physical Sciences Research Council (EPSRC); Ministry of Defence (MoD); Defence Science and Technology Laboratory (Dstl); Centre for the Protection of National Infrastructure (CPNI).

Who has reviewed the trial?

This trial has been reviewed and given favourable opinion by the Ministry of Defence Research Ethics Committee (MoDREC).

Further information and contact details.

For further information, please contact the principal investigator for the trial, Prof. Shaun Lawson, at: shaun.lawson@northumbria.ac.uk

Compliance with the Declaration of Helsinki

This trial complies, and at all times will comply, with the Declaration of Helsinki as adopted at the 64th WMA General Assembly at Fortaleza, Brazil in October 2013.

Appendix B

The Moral Compass Workshop Schedules

B.1 Workshop 1

This document details the running order of the workshop and what is happening at each stage. Workshops should run to approx 60 minutes.

1. Introductions by all members of the workshop - [15 mins]

    - “Hi my name is..." “we’re researchers at Northumbria/Bath"
    - Thanks for coming along, really appreciate you taking up your time
    - “The workshop today will take about 60 minutes, but might run slightly longer. Most important thing to say is that this will be a relaxed session, there’s no tests or trick questions, we really want to have a bit of discussion and carry out some simple tasks"
    - “To kick off, it would be great to go around the room and everybody introduce themselves. I’ll start, my name is [name] I am [job]"
    - Give out pieces of card and pens for people to write their name on and put in front of them on the desk - only do this if sitting around a desk
    - We have pens and paper for you to write on, and please help yourself to food and drink!

2. Introduction to context of the research and brief summary of what we'll do in the workshops - 10 minutes - what will be expected in each workshop and homework

    - Chat about this research and why it's interesting - talk about "Othering" and stigmatisation of people online.
    - This is a totally natural part of human behaviour and we do it all the time - it's important to help us build our cultural and societal identity. For example I am from London and those people there are not.
    - However it can also be used negatively, to high how different a group of people are. For example, *they* are Travellers and don't abide by *our* rules.
    - Another, more extreme example, is TV coverage of ISIS in the middle east, they are referred to as "medieval" - this implies they are not like us (we are not medieval) - this is negative Othering.
    - Talk about the structure of the three workshops
    - The homework
    - And the Amazon vouchers

3. Semi-structured discussion about experiences of using social media. [15 to 20 minutes]

    - Explore their assumptions and experiences of Othering and stigmatisation on social media
    - What are people's experience of social media - what do people use - i.e. Facebook, Instagram, twitter
    - What do people think are interesting events that can be followed on social media (Twitter), and what sort of things do they think are tagged that are examples?

4. Technology

    - Check people's phone types - and can they write down their model - we suggest installing Chrome - but that's optional

- Check if people have laptops and what are they - same again - will they install Chrome - optional

5. Introduction to Moral Compass
    - Hand out how-to sheets
    - Make sure everyone is able to get it working
    - Ensure people have sheets with address and the passwords
6. Example coding session
    - Using projector, show series of tweets and ask the group how they would tag them
    - Using MC coder - start STOMP using a pre-defined (busy) hashtag, get them to do some coding of their own.

Homework 1 Detail

Summarise two live broadcast TV events. These will be events that have active Twitter streams, participants will be asked to use the MC Coder to give morality to tweets in those streams.

Check the times that people can make - suggest they open browser between Wednesday 9pm / Friday 9pm - for an hour and we will pick something fun - maybe that comes out of their suggestions e.g. US Celebrity Big Brother.

B.2 Workshop 2

Participants invited back to discuss how they got on with the tagging exercise. They will also be introduced to the Moral Compass interface, and will give a live broadcast to simulate how tagging and exploring work together.

1. Introduction - find your tags

    Tags from each user will be printed out, as individual “tag clouds" (sensitising material). First task we will ask participants to go look at the tag clouds, figure out which one they think is theirs - when they’ve found it sit down with it.

[Keep a list of whose is whose, in case they don't recognise their own]

2. Experience of tagging

    - Overall how did you get on with the tagging?
    - Give each participant their top *20* or so tweets (sensitising material).
    - Was it easy? Hard? How much stuff did you ignore, why?
    - Were there any tweets where you were torn between tags? Did you give things multiple tags?
    - How did the tagging process make you feel? Angry? Happy? Vindicated?

3. Focusing on the tags themselves:

    - What did you use it for? Did you use it more than once?
    - Why was it important to call it that?

4. Give out and explore "deck" of tweets (sensitising material)

    - Each participant has a "deck" of their own tagged tweets.
    - Did you go to the link?
    - If they want, they can dig through and find a good example
    - We can use the compass to look at their tweets, maybe prompt them on interesting tags if conversations drops.
    - How would you feel if you were seeing all the tagged tweets that you created? Instead of an "untagged" twitter stream, how do you think you'd react to your newly created tagged twitter stream?
        - Would it affect the way you used Twitter?
        - Would you like it? If it was you doing the tagging for yourself (if that could work)?

5. Compass - introduction to how it works

    - Get compass on screen, get participants to load it up on their phone - logging in with their username

- Demonstrate how to scroll on the compass, and that it responds to a tagger who is in the background

6. Pairing participants (tagger and explorer) - logistical issues.
    - Who has access to live UK TV?
    - Will pair randomly, but practically

Homework 2 Detail

Give postcard to record their experience, with 3 questions:

- What did you like the most?
- What did you like the least?
- What would you change?

Britain's Benefit Tenants - Wednesday 9pm - Channel 4
We will email you to tell you what you will be doing (Tagging or Exploring). Please check your email.

B.3 Workshop 3

1. Envelopes to each participant - inside is their username
    - Ask first who they think their pair was - why?
    - Reveal their username from the envelope - go and sit with the other person from pair.
2. Pair reversal activity [10 mins]

    Seat participants on tables in their pairs, audio record each pair. Researcher sitting with each one, prompting them. Get explorers to ask "why did you tag it that way" if they disagreed, or are curios.
    - Those using the explorer: Given a set of printed tweets, given a set of printed tags, asked to recreate their compass (matching tags and tweets)

- Do you remember any of these?
- Did you think there were this many/few?

- Those using tagger: Given the compass to explore digitally, see how their tags looked.
  - What do you think of the theme of your tags? Happy? Critical? Angry?
  - What do you think about the order of the tags?
- Bring everyone back together for a group discussion.

3. Get everyone to recreate their second-screening experience
   - Each person: what were you doing? Who were you with? What device were you using?
   - Use postcards as a jumping off point - probe them on their responses
     - What did you like? Like least? Want to change?
     - Explorers:
       * Did it add anything to the TV viewing experience?
       * What was your “go to" tag?
       * What did you do during the ad breaks?
     - Taggers:
       * Were you waiting for tags/tweets to appear? Were you bored whilst you were waiting?
       * What did you do during the ad breaks?
       * Did you stick on a tag you liked? E.g. hilarious / critical
       * Were you waiting for tags/tweets to appear? Were you bored whilst you were waiting?
     - Final discussion
       * Is it a good or bad tool for highlighting ‘Othering’? Did the tagging/exploring highlight these? If someone disagreed with your tag, would that prompt them to think about it?

* Did anything you came across make you think differently about what you were seeing on-screen? Or did it support what you already thought?
* How would you feel if the tagging process was being done to your own Twitter stream? Select people - How would you feel Y, if X tagged your tweets? [Only if people feel comfortable]
* How about if you could “tune in" to different taggers and see what they thought of a Twitter stream. Were there any difficult aspects to the tagging or exploring, as in things that might have frustrated you (both with the app as well as the general idea of tagging/exploring tweets in this way)? Any things you might want to see incorporated into the app?
* Is the way tweets are presented within the app better or worse than how they are presented on Twitter (i.e. linear/top to bottom)? Both in terms of ease of navigating around them as well as as a user interface.
* Was it easy to use the app alongside your TV viewing or did it get in the way?
* What kinds of TV programme might you prefer/alternatively want to use the app for (not necessarily to do with othering)?

Appendix C

Spotting Guide Participant Information Sheet

Study title

CuRAtOR: Challenging online feaR And OtheRing: Sandbox Trials (Study One)

Invitation to take part

Thank-you for expressing an interest in this research project.

You are invited to participate in a research study being conducted across the Universities of Northumbria, Newcastle, Bath, Aberdeen and Nottingham. This study makes up an element of a wider research initiative named CuRAtOR (Challenging online feaR And OtheRing).

Before deciding to take part in the study, we encourage you to read the following information which will outline why the research is being conducted and what you participation would involve. Please take a period of at least 24 hours to consider whether or not you would like to be a participant in the study - if you are happy to be involved, please reply to the email address provided with a copy of your questionnaire.

What is the purpose of the research?

The aim of CuRAtOR is to explore how software applications might be developed to positively counteract 'Othering' - the practice of making distinctions between groups

of people, often negatively - in online environments. To this aim, the present study aims:

1. To ascertain how people use existing digital applications and services.
2. To ask people to provide their thoughts on an emerging suite of prototype CuRAtOR software applications.

Who is doing this research?

This research will be conducted by various members of the CuRAtOR research team, and supervised by the CuRAtOR principal investigator, Prof. Shaun Lawson. All members of the research team will work to ensure that the research abides by professional ethical standards. Why have I been invited to take part? You have been invited to take part in the study because you have expressed an interest as a result of our advertising this study, and you have identified yourself as someone who fits our eligibility criteria:

- Aged over 18.
- Comfortable with the usage of a one or more social media platforms and other online applications and services (for instance: Twitter, Facebook, YouTube, Reddit, online multiplayer games such as World of Warcraft, etc).
- Have technology capable of running one or more CuRAtOR-developed software applications (i.e. laptop or desktop PC and/or internet-enabled smartphone).

Do I have to take part?

Participation in this study is on a strictly voluntary basis - you do not have to take part in the study and can withdraw from the study at any point, without providing a reason, should you choose to do so.

What will I be asked to do?

Firstly, if you are interested in participating in the study, we request that you fill in and return the attached questionnaire, which will help us understand more about the backgrounds of people taking part in the study. Participants will be selected so as to ensure that people from a wide range of backgrounds are able to contribute.

Should you be selected for the study, we will ask you to sign and return a form consenting to your participation. Once we have selected filled the study with a cohort of volunteers, we will unfortunately have to inform (via email) any additional participants that we are unable to include them in the study.

If you are selected, you will be asked to come to the University of Northumbria to participate in three 'focus groups' (i.e. groups of between 2 and 5 participants). The first focus group will involve a discussion of how you use and think of a selection of existing 'everyday' software applications (i.e. web sites, social media platforms, social games, etc) in terms of what you think about them and how you may already use them yourself, as well as how they may contribute towards online 'Othering'. This focus group will take approximately 60 minutes.

Upon completion of this first element of the study you will be asked again to come to the University of Northumbria to participate in a second focus group (which again will feature between 2 and 5 participants). This focus group will involve the research team demonstrating a selection of prototype software applications developed by the CuRAtOR team, which will form the basis of a participant-led discussion of them. We are particularly interested in finding out which of our technologies you think are most effective (and why), and how the software might be further improved. This focus group will take approximately 60 minutes. The third focus group will be to discuss your use of the prototype software during your homework exercise, as well as your reflections on all of the focus groups. This will take approximately 60 minutes.

You will be asked to perform a homework task between focus groups one and two and focus groups two and three.

You are free to participate in any and all of the elements of the outlined study, but are not obligated to do (i.e. you retain the right to drop out of the study at any time, without providing a reason). Furthermore, if you are willing, you will have an opportunity to take part in a second study which follows on from the study outlined

here. Details on what Study Two involves will be made available in due course, and are available on request from the Principal Investigator, Prof. Shaun Lawson (shaun.lawson@northumbria.ac.uk).

Details of dates and times of the focus group meetings will be provided in due course.

What are the benefits of taking part?

There are no direct benefits for those individuals taking part in this study.

What are the possible disadvantages and risks of taking part?

The study outlined here has been designed such that we do not anticipate any risks involved that would go beyond those experienced in everyday life. However, the software under discussion will at times feature socio-political content which some participants may disagree with. In the remote eventuality that participants become upset upon encountering this content, we will provide information about persons or groups who may be contacted that can help support you in the event you feel distressed. Furthermore, we remind participants that they are free to withdraw from the study at any time and without providing a reason.

Can I withdraw from the research & what will happen if I don't want to carry on?

Yes, you are free to withdraw from the research at any point - you do not need to provide a reason for doing so, and there are no penalties should you choose to withdraw.

Furthermore, you are entitled to request to withdraw any data you do provide for the study, for up to 30 days after completion of the study. After this point, you data will have been anonymised and as such, it will no longer be possible to identify you and remove your data.

Are there any expenses and payments which I will get?

Upon completion of the study, participants will receive a £40 high-street voucher of their choice as a token of appreciation for taking part in the study. Also, any reasonable travel expenses you may accrue through participation in the study will be reimbursed. We will provide information as to the process of how to submit your travel expenses throughout the course of the study.

Whom do I contact if I have any questions or a complaint?

Any questions may be addressed to the principal investigator for the study, Professor Shaun Lawson, who can be contacted at: Dept. of Computer Science and Digital Technologies University of Northumbria Newcastle-upon-Tyne, NE1 8ST Telephone: 01522 886318 E-mail: shaun.lawson@northumbria.ac.uk

If you have a concern or complaint and wish to speak to a person who is independent from the study, you can contact the Head of Computer Science at the University of Northumbria :- Prof John Woodward, Associate Dean (Research and Innovation), Faculty of Engineering & Environment, University of Northumbria, Newcastle, NE1 8ST Telephone: 0191 227 3048 E-mail: john.woodward@northumbria.ac.uk

What happens if I suffer any harm?

We do not anticipate that participants will suffer any harm as a result of their participation in the study. However, if you have any concerns about this possibility, we encourage you to get in touch with the research or supervisory team with the details provided, all of whom will be able to ensure you receive the proper advice and support. In the unlikely event that you suffer any harm from participating in this research, the University of Northumbria will support affected participants in the claiming of compensation in accordance with the Ministry of Defence No Fault Compensation Scheme (for UK nationals) or the University of Northumbria's indemnity policies (for non-UK nationals) - in either eventuality, participants are advised to contact a member of the research team for further advice.

Will my records be kept confidential?

Any and all data collected through the undertaking of this study will be stored confidentially and used exclusively for research purposes (in line with professional ethical standards). The way the data is collected and stored will ensure that individual participants are not identifiable from the data, and any material we use from this data will be thoroughly anonymised. Participants are free to contact the research team at any point during the project or up to 30 days after it has finished to request that their data is not used. In this eventuality, we will immediately destroy all records of the information that you have provided. However, after 30 days your data will have been anonymised and as such we will no longer be able to identify you to exclude your contribution to the study.

Who is organising and funding the research?

This work forms part of the larger project CuRAtOR: Challenging online feaR And OtheRing which was funded through the Empathy and Trust In Communicating Online (EMoTICON) Sandpit initiated by the Economic and Social Research Council (ESRC), in partnership with Arts & Humanities Research Council (AHRC), Engineering and Physical Sciences Research Council (EPSRC), Defence Science and Technology Laboratory (Dstl) and Centre for the Protection of National Infrastructure (CPNI). This work operates under the Ministry of Defence Research Ethics Committee (MoDREC) Protocol Number: 639/MoDREC/15.

Who has reviewed the study?

This study has been reviewed and given favourable opinion by the Ministry of Defence Research Ethics Committee (MoDREC). Further information and contact details. For further information, please contact the principal investigator for the study, Prof. Shaun Lawson, at: shaun.lawson@northumbria.ac.uk

Compliance with the Declaration of Helsinki

This study complies, and at all times will comply, with the Declaration of Helsinki as adopted at the 64th WMA General Assembly at Fortaleza, Brazil in October 2013.

Appendix D

Spotting Guide Workshop Schedules

D.1 Workshop 1

This document details the running order of the workshop and what is happening at each stage. Workshop will run approximately 60 minutes.

1. Introductions by all members of the workshop

    - “Hi my name is..." “we’re researchers at Northumbria"
    - Thanks for coming along, really appreciate you taking up your time
    - The workshop today will take about 60 minutes, but might run slightly longer. Most important thing to say is that this will be a relaxed session, there’s no tests or trick questions, we really want to have a bit of discussion and carry out some simple tasks
    - “To kick off, it would be great to go around the room and everybody introduce themselves. I’ll start, my name is [name] I am a [job]"
    - We have pens and paper for you to write on, and please help yourself to food and drink!

2. Brief summary of what we will do in the workshop

    - Introduce this research and why I am doing it
    - Take a look at and play with an app I’ve made

- A homework task.
- As mentioned this is the first of three workshops.

3. Introduction to Othering, and the context of this research, briefly

   As you will have seen on the poster/email we are interested in “Othering", in particular we’re interested in how traditional media like TV can be used to stereotype and portray people in a bad way.

   We’re technologists as well as researchers, so we are looking for ways we can make technology to help people reflect on these portrayals. This is why we are running the workshops today.

   Othering is the process of creating a distinction between two groups, the familiar group (the *self*) and an outside group (the *other*).

   This is a totally natural part of human behaviour and we do it all the time - it’s important to help us build our cultural and societal identity. For example I am from London and those people there are not.

   However it can also be used negatively, to highlight how different a group of people are. For example, *they* are Travellers and don’t abide by *our* rules.

   Another, more extreme example, is TV coverage of ISIS in the middle east, they are referred to as “medieval" - this implies they are not like us (we are not medieval) - this is negative Othering.

4. Othering on TV

   A freeform discussion - get people talking about portrayal of people on TV.

   - TV shows such as reality TV can make things more dramatic or look worse than they actually are. Can anyone think of any examples?
   - Are there certain types of show which make people look worse?
   - What about “poverty porn" like Benefits Britain, Benefits Street etc?
     - Has anybody ever seen those shows?
     - Do you think they represent all people experiencing poverty in the UK?

- Are there particular things the TV producers do to make things more dramatic?
  - Did anyone see The Great British Bake-Off where that guy put his cake in the bin? Do you think they represented that fairly?
  - Why do you think they do this?
- Do you think TV shows should represent people fairly?
- When, if ever, is it acceptable to negatively other people?
  - ISIS? Terrorists?
- Are there any examples of programmes that positively Other/portray people?
  - Is this a good thing?

5. Verbal Spotting Guide activity, using two clips

   Ok, so the idea now is maybe to see if we can spot any Othering or negative portrayals of people in some TV clips. What we'll do is watch two clips of TV shows, and have a chat about each clip afterwards.

   Clip 1: Benefits Street Series 1 Trailer

   [Discussion of the clip]

   Clip 2: You Can't Get The Staff, Trailer

   [Discussion of the clip]

6. Spotting Guide app introduction

   I built a web app that allows you to "spot" different features of TV shows, with the goal of getting you to think about the way a TV show's producer is crafting and positioning people and stuff (i.e. subtly Othering people).

   We want to explore whether you can actually use it to spot these kinds of things.

   The app, called "The Spotting Guide" is like one of those Michelin I-Spy guides for children you would get when you go on holiday, which say stuff like "Look out for yellow cars" and you mark it down when you spot it.

The idea is you sit down in front of a TV programme, such as Benefits Street, you hit start on the Spotting Guide app, and then you spot the items as you see them.

If you see a common pattern that isn't in the Spotting Guide, you should add it yourself at the bottom.

Just like we did with the last two clips, but using the phone to record things when you see them.

[Show the app very quickly on tablet, or screen shots] Be sure to demonstrate:

- Login (names, passwords)
- How to press play
- How to spot things
- How to add new categories
- How to finish

7. App test with everyone using it on their devices

   Give out sheets that show participants username and the link (with QR code) to the app.

   Users must log in using Safari (Apple) or Chrome (Android). Should work on Windows phone with Chrome.

   Participants MUST use the number they have been given (p1, p2 etc) to login

   Trouble shoot problems people are having with phones etc.

8. Watch a clip of "Big Fat Gypsy Wedding" and do spotting on the app.

   We will delete all this data, so you can experiment with using the app, doing loads of spots, how to pause it and stuff.

   [Alternate video to view if problem with original video: "Nightmare Tenants, Slum Landlords"]

   Reiterate all data will be deleted after this, it is a practice.

9. Homework task

   As mentioned earlier on, we would like you to do some spotting on the Spotting Guide in your own time. We would like you to go away and watch at least 20 minutes of Benefits Street, using the Spotting Guide to spot behaviours.

   Can watch the show a computer or laptop - you need to keep your phone free for using the Spotting Guide. Is this going to be a problem for anyone?

   We need you to do at least 20 minutes of spotting on Benefits Street, starting from the beginning, but you are welcome to do the whole thing. We will use the data you generated in the next workshop!

10. Final Housekeeping

    Thanks again for coming along. We have a £10 Amazon voucher here for each of you today, as compensation for your time.

    We will have a further £30 in Amazon vouchers for each of you, which we will distribute in the final workshop.

    To recap - if you can go away, keep hold of the info sheets we've given you, and watch at least 20 minutes of the Benefits Street show we have linked you to.

    Any questions?

D.2 Workshop 2

This document details the running order of the workshop and what is happening at each stage. Workshop should run for 60 minutes.

Participants will have gone home and done homework after Workshop 1, watching episode 1 of Benefits Street for at least 20 minutes, whilst using the Spotting Guide to spot stuff. They have been shown how to use it and given a how-to sheet, and told to make new categories as they like.

1. Short discussion (around 25 mins) of how participants got on with the homework, some prompts below:

- Did everyone use the app? If not why not?
- How much of Benefits Street did everyone watch?
- Some thoughts on using the app?
- How did it affect the way you watched the TV show?
  - Did you concentrate more on one than the other?
  - Have you “second-screened" before? i.e. using Twitter whilst watching an on-air show?
- How did you get on with the pre-made categories?
- Did you make new categories? A lot of them or a few?

2. Focused discussion, using stats generated from Spotting Guide data (sensitising material). There are three types of data generated: aggregated spots over time, list of all categories, noting new ones, individual graphs of spots over time.

   Discussion can be based around these visualisations:

   (a) List of categories noting new ones:
   - How did everyone get on with the default categories?
   - Who made this category? What made you think of that category?
   - Which categories do you think are “Othering" categories?

   (b) Aggregated spots over time:
   - Why do you think are there spikes?

   (c) Individual graphs:
   - Discuss patterns and features of data.

3. Homework: We’d like you to pick your own show that you think would be interesting to use the spotting guide on. It doesn’t have to be as overt as Benefits Street, you might be interested in looking at the “subtle" stuff in a TV programme.

- If you could watch at least 20 minutes of your selected show, and do some spotting with the spotting guide. We are going to give you a “blank" Spotting Guide, meaning you’ll have to make the categories yourself.
- Next week we’ll all get back together and talk about the shows you picked and what you found

D.3 Workshop 3

This document details the running order of the workshop and what is happening at each stage. Workshop may run to 60 minutes, but could be shorter.

The participants will have gone off to do their homework, which involves selecting a show of their own choice and doing spotting on it (for at least 20 minutes). They were not provided with any spotting categories so they will have had to create their own categories.

1. Check everyone has done the spotting
2. Go around the group and “present":
    - What you watched
    - Why you picked it
    - What kinds of things did you spot on the spotting guide
    - What things did you spot that you found interesting or surprising?
3. Short reflective session about the spotting guide itself
    - Did it change your behaviour when watching the show?
    - Touch on previous points: have you done this before? How would you use it in future?

Appendix E

Screenr Participant Information Sheet

Trial title

CuRAtOR Smartphone App Trial: Viewing and Critiquing TV Production.

Invitation to take part

Thank-you for expressing an interest in this research project.

You are invited to participate in a research trial being conducted by Northumbria University, in Newcastle-upon-Tyne. This trial makes up an element of a wider research initiative named CuRAtOR (Challenging online feaR And OtheRing).

Before deciding to take part in the trial, we encourage you to read the following information which will outline why the research is being conducted and what participation would involve.

What is the purpose of the research?

The aim of CuRAtOR is to explore how software applications might be developed to positively counteract 'Othering' - the practice of making distinctions between groups of people, often negatively - in online environments. As an example, refugees are commonly 'othered' through their depiction as criminals, or a threat to national identity, casting them as 'other' or different from the main population.

In this trial, we are particularly interested in ‘Othering’ that might occur around live TV broadcasts and their associated social media, with our aim being:

- To ascertain how people can decide together what sort of live TV broadcasts might cause or be associated with the practices of ‘Othering’.
- To ask people to use new software applications to arrange to watch live TV broadcasts at the same time from their respective homes.
- To look at new software applications (so called second screen software applications) which allow people to reflect and critique live TV broadcasts during their transmission.

Who is doing this research?

This research will be conducted by various members of the CuRAtOR research team, and supervised by the CuRAtOR principal investigator, Prof. Shaun Lawson. All members of the research team will work to ensure that the research abides by professional ethical standards.

Why have I been invited to take part?

You have been invited to take part in the trial because you have expressed an interest because of our advertising this trial, and you have identified yourself as someone who fits our eligibility criteria:

- Aged over 18.
- Comfortable with the usage of a one or more social media platforms and other online applications and services (for instance: Twitter, Facebook, and YouTube).
- Regularly watch “prime time" TV
- Have technology capable of running a CuRAtOR-developed software application (i.e. internet-enabled smartphone).

Do I have to take part?

No. Participation in this trial is on a strictly voluntary basis - you do not have to take part in the trial and can withdraw from the trial at any point, without providing a reason, should you choose to do so.

What will I be asked to do?

You will be asked to use a smartphone application (app) whilst watching TV to look for specific patterns or techniques used by the programme makers, and record these on the smartphone app (somewhat similar to bingo). The trial lasts for 4 weeks, and we will ask that you watch one programme per week (30/60 minutes) whilst using the app. Each week, along with all the other participants, you will look at a TV guide inside the app (selected by us), and vote which programmes you would like to watch. In the third week of the trial, a member of the research team will contact you via telephone to talk through your experiences so far and troubleshoot any problems. In the fourth week of the trial, you will be requested to visit Northumbria University to attend a focus group to discuss your overall experiences using the app. Each week will commence with voting, and then the app will change to live TV viewing mode for the voted programme. This cycle will repeat each week until the end of the trial.

When voting, you will be able to discuss, using the app, which programmes from the TV guide you think would be interesting to watch and critique. You will be required to cast a vote for your preferred programme. We anticipate this process will be relatively short (approx. 15 minutes), but will vary depending how much you contribute to the discussion with other participants about the programmes before voting.

When watching live TV, the app will become a list of things to look out for in the specific programme. There will be a wide variety of different types of programme being watched because of the voting system, such as reality TV, documentaries, and dramas. When the programme starts, you will be asked to look out for the things in the list, and press on them when you see them. This is somewhat similar to the game bingo, although you are not looking to “win", but to see the things in the programme. For example, if the programme is reality TV focused on welfare claimants, the list will contain things that are commonly focused on in these types

of programmes, as well as techniques the programme makers use e.g.: “Wearing a dressing gown in middle of the day", “Urban deprivation" and “Editing out parts of conversation". Each time you see or notice these things you would record them on your app.

You may express interest by email, replying to the email sent to you by the research team, containing this document. Should you be interested in participating in the trial, we will arrange an initial meeting (at Northumbria University) where we will introduce you to the trial and its aims, and ask you to read and sign a consent form for your participation in the trial. We will then set up the app on your smartphone and provide information on how the trial will proceed. During the trial, we will then need to contact you via telephone on two further occasions so we can get feedback from you on the experience. These telephone interviews will each last approximately 15 minutes, and we will arrange the times and dates of these with you. At the end of the trial, you will be asked to come to Northumbria University to participate in a focus group. This focus group will give the researchers opportunity to find out what you thought of the experience with 4 other participants in the trial - topics will cover how participants used the application for talking together online, scheduling a TV programme to watch and watching this programme at the same time as others, and interacting with others over the Internet.

You are invited to participate in all the elements of the outlined study, but you retain the right to drop out of the trial at any time, without providing a reason. Details of dates and times of the activity schedule will be released in due course.

What is the device or procedure that is being tested?

This trial will test a smartphone application (app) developed by the CuRAtOR team, that will explore how people understand where and how different visual elements and television production techniques are used, as well as understand how you, the participants, recognise and understand these elements. The app will work in two phases:

1. a voting system for all participants to look at a set of TV listings that we have compiled, with you all discussing and voting for the programmes you think would be most interesting to watch and critique.

2. a live TV viewing phase, where you will look out for specific things in the programme, and record them when you see them. For example, one item in the list may be “Dramatic Music", and you will be asked to record that you’ve spotted Dramatic Music in the programme, by pressing on the item in the list.

You will be provided with a username and password, uniquely assigned to you, with which to access the app. We will collect usage information about the smartphone app only, for example when the app is started and stopped, how and when votes are cast, and when specific items are recorded during TV viewing. This is to help us understand how the app is being used, as well as to tailor our interviews and focus groups to the participants. This data will be fully anonymised and you will not be identifiable in any subsequent publications. All usage information will be transmitted securely to our servers, and this information will not be stored on your phone at any point. At the conclusion of the trial, the research team will demonstrate how to permanently uninstall the app from your phone.

What are the benefits of taking part?

There are no direct benefits for those individuals taking part in this trial.

What are the possible disadvantages and risks of taking part?

The trial outlined here has been designed such that we do not anticipate any risks involved that would go beyond those experienced in everyday life. However, the software under discussion will at times feature socio-political content which some participants may disagree with. In the remote eventuality that participants you become upset upon encountering this content, we will provide information about persons or groups who may be contacted that can help support you in the event you feel distressed. You should be aware that we adhere to The British Psychological Society Ethics and Code of Conduct, and therefore if you disclose anything that raises concerns relating to duty of care, the researcher has a responsibility to act on this information. If this occurs, the researcher will advise you at that time, and reiterate the implications of continuing to reveal such insight (i.e. the loss of guarantee of

confidentiality) and ask you to consider if you want to continue that discussion. Furthermore, we remind participants that they are free to withdraw from the trial at any time and without providing a reason.

Can I withdraw from the research & what will happen if I don't want to carry on?

Yes, you are free to withdraw from the research at any point - you do not need to provide a reason for doing so, and there are no penalties should you choose to withdraw.

Furthermore, you are entitled to request to withdraw any data you do provide for the trial, for up to 30 days after completion of the trial. Your trial data will be removed, when notified within the 30 day period, however your contribution to the focus group will remain included in the project, due to the nature of audio recording. You retain the right to withdraw from the focus group at any time by ceasing your contribution to the focus group. Beyond 30 days after completion of the trial, all of your data will have been anonymised and as such, it will no longer be possible to identify you and remove your data.

Are there any expenses and payments which I will get?

As a token of our appreciation for taking part, you will receive up to £150 for completing the multiple steps in the trial.

This means that if you are unable to participate in one step of the research for whatever reason, or if you withdraw from the trial before it is completed, you will still receive vouchers for those steps you have already participated in.

Whom do I contact if I have any questions or a complaint?

Any questions may be addressed to the principal investigator for the trial, Professor Shaun Lawson, who can be contacted at:

> Prof. Shaun Lawson, Professor of Social Computing, Computer and Information Sciences, Northumbria University, NE1 8SG, Telephone: +44 (0)191 2273944, Email: shaun.lawson@northumbria.ac.uk

If you have a concern or complaint and wish to speak to a person who is independent from the trial, you can contact the Head of Computer and Information Sciences at the University:

> Graham Sexton, Head of Department, Computer and Information Sciences, Northumbria University, NE1 8SG, Telephone: +44 (0)191 2273232, Email: graham.sexton@northumbria.ac.uk

What happens if I suffer any harm?

We do not anticipate that participants will suffer any harm because of their participation in the trial. However, if you have any concerns about this possibility, we encourage you to get in touch with the research or supervisory team with the details provided, all of whom will be able to ensure you receive the proper advice and support. In the unlikely event that you suffer any harm from participating in this research, the Northumbria University will support affected participants in the claiming of compensation in accordance with the Ministry of Defence No Fault Compensation participants are advised to contact a member of the research team for further advice.

Will my records be kept confidential?

All data collected through the undertaking of this trial will be stored confidentially and used exclusively for research purposes (in line with professional ethical standards). The way the data is collected and stored will ensure that individual participants are not identifiable from the data, and any material we use from this data will be thoroughly anonymised. Participants are free to contact the research team at any point during the project or up to 30 days after it has finished to request that their data is not used. In this eventuality, we will immediately destroy all records of the information that you have provided. However, after 30 days your data will have been anonymised and as such we will no longer be able to identify you to exclude your contribution to the trial.

Who is organising and funding the research?

This work is funded by: Economic and Social Research Council (ESRC); Arts & Humanities Research Council (AHRC); Engineering and Physical Sciences Research Council (EPSRC); Ministry of Defence (MoD); Defence Science and Technology Laboratory (Dstl); Centre for the Protection of National Infrastructure (CPNI).

Who has reviewed the trial?

This trial has been reviewed and given favourable opinion by the Ministry of Defence Research Ethics Committee (MoDREC).

Further information and contact details.

For further information, please contact the principal investigator for the trial, Prof. Shaun Lawson, at: shaun.lawson@northumbria.ac.uk

Compliance with the Declaration of Helsinki.

This trial complies, and at all times will comply, with the Declaration of Helsinki as adopted at the 64th WMA General Assembly at Fortaleza, Brazil in October 2013.

Appendix F

Screenr Interview and Focus Group Schedules

F.1 Screenr Entrance Interview

The Week 1 Interview will be introductory in nature, with the overall goal to meet participants and explain/answer questions about the nature of the trial, show them the software application they will be using over the course of the trial, and instruct them in its features and functions. Also, to briefly interview the participant about their views and experiences of existing technologies relevant to second-screening. Introductions to the research from the researcher The sessions will begin with a brief introduction from the members of the research team, which will outline the goals of the wider CuRAtOR project and the goals of the present trial.

1. Consent Form.

    - Ensure they have read the consent form and are happy to sign (if they haven't already).
    - Discussion around the types of technologies and platforms participants use
    - Switch on the audio recorder
    - Prompt: Are you familiar with the term othering?
    - Provide an example of othering if they are not.

- Prompt: Do you watch reality TV? If so why do you watch it? If not, why not?
- Prompt: What do you think the TV producers are trying to portray when they make a reality TV programme?
- Prompt: What about the issues raised within these programmes - poverty, state welfare etc?
- Prompt: Do you live tweet or talk about programmes on social media?
- Prompt: What types of digital platforms and technologies do you use when watching TV?
- Prompt: What is it that attracts you to [specific digital platform/technology]? Why do you use it, and what do you use it for?

2. Setting up and demonstrating the software application The aim here is to ensure the participant will able to access and use the software application on their smartphone. Make sure they are given a copy of the how-to guide.

F.2 Screenr Final Focus Group

Important: By continuing to participate in the focus group they cannot withdraw their data from it (we can't edit them out of the conversation). If they want to withdraw from the focus group they must just leave. However, this does not affect their right to withdraw the other data they have contributed to the study in the form of using the app.

1. *Ask people to introduce themselves (for the audio)*

   They do not have to share their username in the focus group

   Experiences using the software application, and of the software design

2. Prompt: Was the software application easy to use? Was it easy to understand how to navigate the software? Was it easy to see how the various functions of the software application worked and what they did?

- Prompt: Was it easy or difficult to use the software alongside watching a TV programme? Were there aspects of the software application that were difficult or frustrating to use?
- Prompt: If you were designing the software application yourself, what would you change? Is there anything you might include that we haven't already? Anything that you would take out that is already in there? Thinking about the experience of use - overall
- Prompt: How did the voting mechanism play out? Did you make your decisions immediately? Did you share the programme ideas with other people?
- Did you get to watch what you wanted to watch? What you voted for? How did it make you feel to not get the programme you wanted?
- Did you agree with everyone's comments? Did you see things you disagreed with? Where did you see these mainly (chat/tags)
- When using the app whilst watching TV, where did you spend most of your time? Programme choices - refer to programme cards
- What did you vote for? Why? What didn't you vote for?
- Did scheduling/time play a factor? Did the content or critical aspect come into it?
- What would have been your "ideal" programme?

3. Weekly Word Clouds - refer to weekly word clouds (sensitising material)

    - Do you think the tags were useful to record what was going on in the programme?
    - Were the tags useful to critique the programme?
    - What do you think about the content of the weekly tags? Any surprising or interesting things there?
    - Would the weekly tags clouds tell you anything if you hadn't seen the programme?

4. Concluding remarks by researchers and session conclusion. Wrap up the session, thank everyone for their participation. Ask if anyone has any questions about the project, we can take email addresses if participants would like to know where the research leads.

Appendix G

Spkr Participant Information Sheet

Trial title

CuRAtOR Device Trial: Interacting with Twitter content through a talkative home device.

Invitation to take part

Thank-you for expressing an interest in this research project.

You are invited to participate in a research trial being conducted by Northumbria University, in Newcastle upon Tyne. This trial makes up an element of a wider research initiative named CuRAtOR (Challenging online feaR And OtheRing).

Before deciding to take part in the trial, we encourage you to read the following information which will outline why the research is being conducted and what participation would involve. Please take a period of at least 24 hours to consider whether you would like to be a participant in the trial.

What is the purpose of the research?

The aim of CuRAtOR is to explore how software applications might be developed to positively counteract 'Othering' - the practice of making distinctions between groups of people, often negatively - in online environments. As an example, refugees are commonly 'othered' through their depiction as criminals, or a threat to national identity, casting them as 'other' or different from the main population.

In this trial, we are particularly interested in ways of interacting with discussion from the Twitter platform, with our aim being:

1. To explore people's interaction with different types of online discussion and statements, and how they might respond.
2. To understand how people interact with a home device providing them with spoken piece of Twitter content in the home environment.

Who is doing this research?

This research will be conducted by various members of the CuRAtOR research team, and supervised by the CuRAtOR principal investigator, Prof. Shaun Lawson. All members of the research team will work to ensure that the research abides by professional ethical standards.

Why have I been invited to take part?

You have been invited to take part in the trial because you have expressed an interest because of our advertising this trial, and you have identified yourself as someone who fits our eligibility criteria:

- Aged over 18.
- Do not have a Google Home, Amazon Alexa or similar smart home device within your house
- Have a home Wi-Fi network
- Are not away from home for more than 5 days consecutively for the duration of the study.

Do I have to take part?

No. Participation in this trial is on a strictly voluntary basis - you do not have to take part in the trial and can withdraw from the trial at any point, without providing a reason, should you choose to do so.

What will I be asked to do?

You will be asked to interact with a home device, called Spkr, that will “speak" pieces of discussion from the online social media site Twitter. You will be asked to listen to these broadcasts, and reflect on the content (whether you agree, disagree, have something to say), and should you so desire respond to the device by saying a specific phrase, followed by your thoughts.

The Spkr device will be installed in your house for the period of 28 days, where it will speak 10 broadcasts per day. When we come to install Spkr in your house we will ask you to specify three time periods where you would like Spkr to be switched on. For example, this might be 8am until 10am, 6pm until 8pm, 9pm until 10pm. Spkr will then only play within those time periods for the duration of the 28 days. These times will be the same for each day of the week.

In the first week of the trial, we will schedule a time with you for two members of the research team to come to your house. We will install Spkr in your desired location, as well as conduct a short 20 minute interview with you. Spkr will then start speaking on the following day, which is classed as day 1 of the study.

We will contact you on day 3 to check all is ok with the device and answer any questions you may have. Around day 14 we will contact you again to ensure all is working correctly.

After day 28, Spkr will stop speaking, and we will again schedule a time with you for two members of the research team to come to your house and take Spkr away. We will then conduct a short 20 minute interview with you about your experiences.

During the entrance interview, we will ask you to fill out a political alignment questionnaire. This is to allow us to align the content being spoken to you through Spkr to your political alignment.

All content spoken by the device is collected from Twitter, with each piece of content being examined by the research team. We will remove swear words and other offensive terms where possible. Much of the content will concern politics or have political tones, examples of what Spkr will broadcast during the day might be:

- *“Don’t you get fed up with having to explain that two plus two equals four rather than five? I don’t get why it’s not obvious to certain people that if someone*

is defrauding the Grenfell funds they must by definition not be victims of the disaster."

- *“Headline: What Theresa May said at Robben Island is an insult. Comment: She’s part of the generation of gross young Tories who called Mandela a terrorist and called for him to be hanged. See former Prime Minister Cameron’s all expenses paid trip to South Africa with an anti-sanctions lobbying firm."*
- *“Sure, let’s give out a huge tax cut, billions more for the military, millions for golf weekends and "love me" rallies, but then "OMG the deficit" and take it out on civilian government workforce"*
- *“I never belived HS2 would reach Leeds by 2033. They only ever needed as a minimum for HS2 to reach Crewe. Then the people who paid politicians who lobbied for HS2 get their rail freight capacity to Liverpool and Salford! It’s the biggest effing corruption of Government in UK history! Hashtag Northern Powerhouse."*

We will ask that you site Spkr in a communal space, such as a dining room or kitchen, so that you hear it speaking. It is ok to miss when Spkr speaks, but we would ask you try to listen to as many as you can.

It is important to note that as Spkr operates in a communal home environment it is possible that others inside the home may hear the content being spoken by Spkr. We will censor all swearing. Given the socio-political nature of the content, you should consider the impact of this on other members of the household. We request that only you interact with Spkr.

You are invited to participate in all the elements of the outlined trial, and you retain the right to withdraw from the trial at any time, without providing a reason. Upon doing so we will be required to collect Spkr from your home.

What is the device that is being tested?

The device being tested, called Spkr, is a home device that “speaks" pieces of discussion from the online social media site Twitter. It is configured to only speak within three time periods each day, specified by you, to avoid unnecessary interruption.

You are able to respond to any of the broadcasts at any time by speaking out loud using a specific phrase. Spkr uses voice recognition to capture what you are saying and send this to the research team.

Spkr will speak at most 10 times per day, each lasting around 30 seconds. You can respond to any of what Spkr says, which is done by saying out loud “Alexa, speaker feedback", at which point the device will say “Please give me your feedback", and will record what you say out loud.

At a technical level, Spkr is made up of two pieces of technology, see image below. Part A is a wooden enclosure that contains a credit-card sized computer, called a Raspberry Pi. The device on top, Part B, is an Amazon Echo (often known as “an Alexa") voice assistant. These two devices work together to create Spkr.

Spkr connects securely to your Wi-Fi. Part A will communicate with our research server, where it downloads the voice audio (what you hear speaking). It will consume approximately 500Mb (0.5Gb) to 1000Mb (1Gb) of broadband usage per month. No other information is transmitted by Part A to our research server.

Data Collection:

We will collect usage information from Spkr. Part A (Raspberry Pi) records when it is switched on, and when and what it speaks each day. This information is stored on the computer and is retrieved at the end of the research trial.

Part B (Amazon Echo) will collect all voice interactions with the it. For clarity, this means when you say the word “Alexa", anything you say subsequently will be captured and stored on Amazon’s server, and will be available to us. For the purposes of this research, we will only collect your voice interactions when you use the Alexa skill “Speaker Feedback", with any other voice content ignored.

We collect this data to help us understand how you are using Spkr, as well as to tailor our interviews to each participant. This data will be fully anonymised and you will not be identifiable in any subsequent publications. Usage information collected by Part A will be stored locally on the device and collected at the end of the trial by the research team. Any voice interactions with Part B (Amazon Echo) will be securely transferred to Amazon’s servers, where only the research team will have access to your voice recordings, which will be anonymised as detailed above. What are the benefits of taking part? There are no direct benefits for those individuals taking part in this trial.

Figure G.1: The two parts that comprise Spkr: Part A: Raspberry Pi computer inside wooden base (bottom). Part B: Amazon Echo speaker (top)

What are the possible disadvantages and risks of taking part?

The trial outlined here has been designed such that we do not anticipate any risks involved that would go beyond those experienced in everyday life. However, the device under discussion will at times feature socio-political content which some participants may disagree with. In the remote eventuality that participants become upset upon

encountering this content, we will provide information about persons or groups who may be contacted that can help support you in the event you feel distressed.

As Spkr operates in a communal home environment it is possible that others inside the home may hear the content being spoken by Spkr. Given the socio-political nature of the content this should be considered.

You should be aware that we adhere to The British Psychological Society Ethics and Code of Conduct, and therefore if you disclose anything that raises concerns relating to duty of care, the researcher has a responsibility to act on this information. If this occurs, the researcher will advise you at that time, and reiterate the implications of continuing to reveal such insight (i.e. the loss of guarantee of confidentiality) and ask you to consider if you want to continue that discussion. Furthermore, we remind participants that they are free to withdraw from the trial at any time and without providing a reason.

Can I withdraw from the research & what will happen if I don't want to carry on?

Yes, you are free to withdraw from the research at any point - you do not need to provide a reason for doing so, and there are no penalties should you choose to withdraw. Furthermore, you are entitled to request to withdraw any data you do provide for the trial, for up to 30 days after completion of the trial. Your trial data will be removed, when notified within the 30 day period, however your contribution to the focus group will remain included in the project, due to the nature of audio recording. Beyond 30 days after completion of the trial, all of your data will have been anonymised and as such, it will no longer be possible to identify you and remove your data.

Are there any expenses and payments which I will get?

As a token of our appreciation for taking part, you will receive £60 in intu Eldon Square vouchers, broken into £10 payable at the beginning of the trial, and £50 payable at the end of the trial.

Whom do I contact if I have any questions or a complaint?

Any questions may be addressed to the principal investigator for the trial, Professor Shaun Lawson, who can be contacted at:

Prof. Shaun Lawson Head of Department & Professor of Social Computing Computer and Information Sciences Northumbria University NE1 8SG Telephone: +44 (0)191 2273944 Email: shaun.lawson@northumbria.ac.uk

If you have a concern or complaint and wish to speak to a person who is independent from the trial, you can contact Professor David Kirk of the Computer and Information Sciences at the University:

Prof. David Kirk Computer and Information Sciences Northumbria University NE1 8SG Telephone: 0191 2273376 Email: david.kirk@northumbria.ac.uk

What happens if I suffer any harm?

We do not anticipate that participants will suffer any harm because of their participation in the trial. However, if you have any concerns about this possibility, we encourage you to get in touch with the research or supervisory team with the details provided, all of whom will be able to ensure you receive the proper advice and support. In the unlikely event that you suffer any harm from participating in this research, Northumbria University will support affected participants in the claiming of compensation in accordance with the Ministry of Defence No Fault Compensation Scheme (for UK nationals) participants are advised to contact a member of the research team for further advice.

Will my records be kept confidential?

All data collected through the undertaking of this trial will be stored confidentially and used exclusively for research purposes (in line with professional ethical standards). The way the data is collected and stored will ensure that individual participants are not identifiable from the data, and any material we use from this data will be thoroughly anonymised. Participants are free to contact the research team at any point during the project or up to 30 days after it has finished to request that their data

is not used. In this eventuality, we will immediately destroy all records of the information that you have provided. However, after 30 days your data will have been anonymised and as such we will no longer be able to identify you to exclude your contribution to the trial.

Who is organising and funding the research?

The Research Project is being led by Northumbria University in collaboration with the University of Bath and the University of Nottingham.

The project is funded by: Economic and Social Research Council (ESRC); Arts & Humanities Research Council (AHRC); Engineering and Physical Sciences Research Council (EPSRC); Ministry of Defence (MoD); Defence Science and Technology Laboratory (Dstl); Centre for the Protection of National Infrastructure (CPNI).

Who has reviewed the trial?

This trial has been reviewed and given favourable opinion by the Ministry of Defence Research Ethics Committee (MoDREC).

Further information and contact details.

For further information, please contact the principal investigator for the trial, Prof. Shaun Lawson, at: shaun.lawson@northumbria.ac.uk

Compliance with the Declaration of Helsinki.

This trial complies, and at all times will comply, with the Declaration of Helsinki1as adopted at the 64th WMA General Assembly at Fortaleza, Brazil in October 2013.

Appendix H

Spkr Participant Entrance & Exit Interview Schedule

H.1 Entrance Interview

Induction Process

1. Consent form
    - Answer any questions that they have
    - Explain the function of the device
2. Political alignment test
    - They will complete this printed. Be on hand to answer queries.
3. What three times would you like the device to be on? Fill in sheet, 3 time slots.
4. Commence interview.
5. Finally - give £10 intu Eldon Square voucher.

Interview Questions

Existing News Consumption Practices

- Would you say you follow the news?

- Where would you say your main source of news information is? Why?
- What format do you consume most? Printed, aural, social media, digital articles?
- (Why / why don't) you get your news from social media?
- Do you feel like the news you consume gives you a balanced view of topics?
  - Do you actively try to get a balanced view?
- Do your views align with those you interact with? (e.g. friends, colleagues, family).

Social Media Use

1. Do you use social media?
   (a) What do you use social media for?
   (b) Do you post stuff yourself?
2. Why don't you use social media?
3. Do you read political discussions on social media?
   (a) Have you come across them? Such as tweets within articles? Do you engage with them?

Engagement in Political Debates On & Offline

1. Do you read the comments on news articles?
   (a) Do you comment yourself?
2. Do you engage in political discussions (in person, or online)? Why/why not?

If online What platforms? Please explain the last two things you discussed?
   (a) How do you handle disagreement with others?

3. How do you think you're aligned politically?
4. Would you say you are politically active? If so, in what ways?

Engagement with smart home technology & audio devices

1. Have you ever had a home device that talked to you? Or do you ever listen to talk radio?
   (a) (examples being Radio 4/LBC, a Furby/Barbie etc, smart home assistant)
   (b) What do you think of them? Did you get any specific use out of them?
   (c) (If a talk radio listener) Do you tend to leave it on in the background? Why?
2. Does anyone in the house bring up news topics for discussion?
   (a) If yes - how does that go down?

Device Setup

1. Setup device, connect to WiFi, etc.
2. Photograph location of device

H.2 Exit Interview

General - warm up

- How did you get on with Spkr over the last month?
- How did you get on with the accent? Were there any ways that you coped with understanding what was being said?
- How often would you say you responded?
- What made you want to respond? Tell me about your motivation to respond

User Experience

- What did you think of having a device that talked to you randomly?
  - When did you notice the device talking most?
- Can you give me an idea of how Spkr fit into your daily life?
  - And into social occasions/when people came over?
  - Lead to awkward situations?
- How did you manage the volume at different times (e.g. social situations etc)?
- How did you explain the device to other people, if they came into the house?
- Can you give me a picture of how other people interacted with it? (It's ok if they did.)

Engagement with Spkr content

- Can you think of any topics that came up that were topics you didn't want to discuss?
- What about the appropriateness of the topics for the home environment?
- How did the topics or content from Spkr feed into household discussion or with friends?
- Can you think of an example where you heard a topic on Spkr and did further research or talked to other people about it?
- What kind of views do you think were represented in the topics?
- Can you give any examples of where you noticed different perspectives in the content?
- How did different perspectives affect the way you looked at news?
- How relevant did you find Spkr content? How informative did you find it?

Changes in engagement with on/offline debates

- How did having Spkr in your house affect your existing news consumption?
- Did you share anything about Spkr or the content on social media?

Changes in news consumption habits / behaviour

- How do you think your awareness of news topics has changed by having Spkr in your house?
- Can you think of an example of a difference or similarity between something you heard on Spkr and the news you normally consume?
- How do you think a system like Spkr (speaking from Twitter, taking discussion) would work in real life?
  - If there was an Alexa skill of Spkr available tomorrow, would you use it?
- How do you think this could be used in the real world? How would topics be chosen every day?
- How do you think your responses could be used? Would you be happy for them to be sent as a reply to the person who sent the tweet? Would you want that shared?

Explore Data Together With Participant

[Handover calendar visualisation displaying the topics per day and number of responses]

Can you think of any topics that really got your interested or engaged?

References

[1] ACM. ACM Code of Ethics and Professional Conduct. *Code of Ethics*, 2018.

[2] Lada A. Adamic and Natalie Glance. The political blogosphere and the 2004 US election: divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, pages 36–43. ACM, 2005.

[3] Philip Agre. Toward a critical technical practice: Lessons learned in trying to reform AI. In Geoffrey Bowker, Susan Leigh Star, Les Gasser, and William Turner, editors, *Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide*. Psychology Press, 1997.

[4] Philip E. Agre. *Computation and human experience*. Cambridge University Press, 1997.

[5] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–36, 2017.

[6] Amazon. Amazon Alexa Skills - news. `https://www.amazon.co.uk/s?k=news&i=alexa-skills`, 2020. [Online; Accessed 27-January-2020].

[7] Morgan Ames and Mor Naaman. Why we tag: motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 971–980. ACM, 2007.

[8] Edward Anstead, Steve Benford, and Robert J. Houghton. Many-screen viewing: evaluating an Olympics companion application. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, pages 103–110. ACM, 2014.

[9] Matthew P. Aylett and Shaun Lawson. The smartphone: A lacanian stain, a tech killer, and an embodiment of radical individualism. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 501–511. ACM, 2016.

[10] Christopher A. Bail, Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37):9216–9221, 2018.

[11] Lynne Baillie and David Benyon. Place and technology in the home. *Computer Supported Cooperative Work (CSCW)*, 17(2-3):227–256, 2008.

[12] Eytan Bakshy, Solomon Messing, and Lada A. Adamic. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239):1130–1132, 2015.

[13] Santosh Basapur, Hiren Mandalia, Shirley Chaysinh, Young Lee, Narayanan Venkitaraman, and Crysta Metcalf. FANFEEDS: evaluation of socially generated information feed on second screen as a TV show companion. In *Proceedings of the 10th European Conference on Interactive TV and Video*, pages 87–96. ACM, 2012.

[14] Eric P. S. Baumer. Reflective informatics: conceptual dimensions for designing technologies of reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 585–594. ACM, 2015.

[15] Eric P. S. Baumer, Claire Cipriani, Mitchell Davis, Gary He, James Kang, Jaclyn Jeffrey-Wilensky, Jinjoo Lee, Justin Zupnick, and Geri K. Gay. Broadening exposure, questioning opinions, and reading patterns with reflext: A computational support for frame reflection. *Journal of Information Technology & Politics*, 11(1):45–63, 2014.

[16] BBC. Telescope 2017 report. `https://www.tvlicensing.co.uk/about/telescope`, 2017. [Online; Accessed 01-January-2017].

[17] Tom W. Bell. The constitution as if consent mattered. *Chap. L. Rev.*, 16:269, 2012.

[18] Erin Beneteau, Ashley Boone, Yuxing Wu, Julie A. Kientz, Jason Yip, and Alexis Hiniker. Parenting with alexa: Exploring the introduction of smart speakers on family dynamics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

[19] Joe Bennett. Chav-spotting in Britain: the representation of social class as private choice. *Social Semiotics*, 23(1):146–162, 2013.

[20] Frank Bentley, Katie Quehl, Jordan Wirfs-Brock, and Melissa Bica. Understanding online news behaviors. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 590. ACM, 2019.

[21] Anol Bhattacherjee. Social science research: Principles, methods, and practices. *USF Tampa Library Open Access Collections*, 2012.

[22] Sam Biddle. Justine Sacco is good at her job, and how I came to peace with her. `https://gawker.com/justine-sacco-is-good-at-her-job-and-how-i-came-to-pea-1653022326`, 2014. [Online; Accessed 10-January-2020].

[23] Jonathan Bignell. *Big Brother: Reality TV in the twenty-first century*. Springer, 2005.

[24] Anita Biressi and Heather Nunn. *Reality TV: Realism and revelation*. Columbia University Press, 2005.

[25] Broadcasters' Audience Research Board. Monthly top 30 programmes. `http://www.barb.co.uk/viewing-data/monthly-top-30`, 2016. [Online; Accessed 21-March-2019].

[26] Broadcasters' Audience Research Board. Monthly viewing summary. `http://www.barb.co.uk/viewing-data/monthly-viewing-summary/`, 2017. [Online; Accessed 29-September-2017].

[27] Pablo J. Boczkowski, Eugenia Mitchelstein, and Mora Matassi. “news comes across when I’m in a moment of leisure": Understanding the practices of incidental news consumption on social media. *New Media & Society*, 20(10):3523–3539, 2018.

[28] David Bodoff and Eran Vaknin. Priming effects and strategic influences in social tagging. *Human–Computer Interaction*, 31(2):133–171, 2016.

[29] Kirsten Boehner, Shay David, Joseph Kaye, and Phoebe Sengers. Critical technical practice as a methodology for values in design. In *CHI 2005 Workshop on quality, values, and choices*, pages 2–7, 2005.

[30] Moira Bovill and Sonia Livingstone. Bedroom culture and the privatization of media use. 2001.

[31] Kevin K. Bowden, Jiaqi Wu, Wen Cui, Juraj Juraska, Vrindavan Harrison, Brian Schwarzmann, Nicholas Santer, Steve Whittaker, and Marilyn Walker. Entertaining and opinionated but too controlling: a large-scale user study of an open domain Alexa prize system. In *Proceedings of the 1st International Conference on Conversational User Interfaces*, pages 1–10, 2019.

[32] Simon Bowen. Crazy Ideas or Creative Probes?: Presenting Critical Artefacts to Stakeholders to Develop Innovative Product Ideas. *Proceedings of EAD07: Dancing with Disorder: Design, Discourse & Disaster*, Izmir, Turkey(April):11–13, 2007.

[33] dana boyd. The power of fear in networked publics. SXSW, Austin, Texa. `http://www.danah.org/papers/talks/2012/SXSW2012.html`, 2012. [Online; Accessed 13-January-2020].

[34] Engin Bozdag and Jeroen van den Hoven. Breaking the filter bubble: democracy and design. *Ethics and Information Technology*, 17(4):249–265, 2015.

[35] Ted Brader. *Campaigning for hearts and minds: How emotional appeals in political ads work*. University of Chicago Press, 2006.

[36] William J Brady, Ana P Gantman, and Jay J Van Bavel. Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General*, 2019.

[37] Virginia Braun and Victoria Clarke. *Successful qualitative research: A practical guide for beginners*. Sage, 2013.

[38] Virginia Braun, Victoria Clarke, Nikki Hayfield, and Gareth Terry. Thematic analysis. *Handbook of Research Methods in Health Social Sciences*, pages 843–860, 2019.

[39] Virginia Braun, Victoria Clarke, and Gareth Terry. Thematic analysis. *Qualitative Psychology: A Practical Guide To Research Methods*, 24:95–114, 2014.

[40] Gabriel Bristow. Yellow fever: populist pangs in France: Reflections on the gilets jaunes movement and the nature of its populism. *Soundings: A journal of politics and culture*, 72(72):65–78, 2019.

[41] Phil Brooker, John Vines, Selina Sutton, Julie Barnett, Tom Feltwell, and Shaun Lawson. Debating poverty porn on twitter: Social media as a place for everyday socio-political talk. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3177–3186. ACM, 2015.

[42] Barry Brown, Alexandra Weilenmann, Donald McMillan, and Airi Lampinen. Five provocations for ethical HCI research. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 852–863. ACM, 2016.

[43] Axel Bruns and Jean E. Burgess. #ausvotes: How twitter covered the 2010 Australian federal election. *Communication, Politics and Culture*, 44(2):37–56, 2011.

[44] Marion Buchenau and Jane Fulton Suri. Experience prototyping. In *Proceedings of the 3rd Conference on Designing Interactive Systems: processes, practices, methods, and techniques*, pages 424–433. ACM, 2000.

[45] Ricardo Buettner and Katharina Buettner. A systematic literature review of Twitter research from a socio-political revolution perspective. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 2206–2215. IEEE, 2016.

[46] Matt Burgess. What is the Internet of Things? WIRED explains. `https://www.wired.co.uk/article/internet-of-things-what-is-explained-iot`, 2018. [Online; Accessed 28-February-2020].

[47] Christopher Buschow, Beate Schneider, and Simon Ueberheide. Tweeting television: Exploring communication activities on twitter while watching TV, 2014.

[48] Michael A. Cacciatore, Dietram A. Scheufele, and Shanto Iyengar. The end of framing as we know it... and the future of media effects. *Mass Communication and Society*, 19(1):7–23, 2016.

[49] Robyn Caplan and danah boyd. Who controls the public sphere in an era of algorithms. *Mediation, Automation, Power*, pages 1–19, 2016.

[50] Nico Carpentier and Bart Cammaerts. Hegemony, Democracy, Agonism and Journalism: An interview with Chantal Mouffe. *Journalism Studies*, 7(6):964–975, 2006.

[51] Manuel Castells. *Communication power*. OUP Oxford, 2013.

[52] Andrea Ceron and Sergio Splendore. 'Cheap Talk'? Second screening and the irrelevance of TV political debates. *Journalism*, 20(8):1108–1123, 2019.

[53] Pablo Cesar, Dick C. A. Bulterman, and A. J. Jansen. Usages of the secondary screen in an interactive television environment: Control, enrich, share, and transfer television content. In *European Conference on Interactive Television*, pages 168–177. Springer, 2008.

[54] Alan Chamberlain, Andy Crabtree, Tom Rodden, Matt Jones, and Yvonne Rogers. Research in the wild: Understanding 'in the wild' approaches to design and development. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 795–796, New York, NY, USA, 2012. ACM.

[55] Despoina Chatzakou, Nicolas Kourtellis, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Athena Vakali. Hate is not binary: Studying abusive behavior of# gamergate on Twitter. In *Proceedings of the 28th ACM conference on hypertext and social media*, pages 65–74, 2017.

[56] Ian Cole. Is a little knowledge about welfare a dangerous thing? A small scale study into attitudes towards, and knowledge about, welfare expenditure. *People, Place & Policy Online*, 9(1), 2015.

[57] Elanor Colleoni, Alessandro Rozza, and Adam Arvidsson. Echo chamber or public sphere? predicting political orientation and measuring political homophily in Twitter using big data. *Journal of communication*, 64(2):317–332, 2014.

[58] Hatty Collier. Channel 4's Benefits Street claims 4.3 million viewers. `https://www.theguardian.com/media/2014/jan/07/tvratings-channel4`, 2014. [Online; Accessed 21-March-2019].

[59] Elena Cresci. Chatbot that overturned 160,000 parking fines now helping refugees claim asylum. `https://www.theguardian.com/technology/2017/mar/06/chatbot-donotpay-refugees-claim-asylum-legal-aid`, 2017. [Online; Accessed 27-January-2020].

[60] Clara Crivellaro, Rob Comber, John Bowers, Peter C. Wright, and Patrick Olivier. A pool of dreams: Facebook, politics and the emergence of a social movement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3573–3582. ACM, 2014.

[61] Jon Crowcroft, Hamed Haddadi, and Tristan Henderson. Responsible research on social networks. In *The Oxford Handbook of Networked Communication*. Oxford University Press, 2017.

[62] Daryl Davis. What do you do when someone just doesn't like you? | daryl davis | tedxcharlottesville. `https://youtu.be/Y4gly9n9RBo`, 2018. [Online; Accessed 27-January-2020].

[63] Simone de Beauvoir. *The second sex*. Vintage, 2012.

[64] Roberta De Michele, Stefano Ferretti, and Marco Furini. On helping broadcasters to promote TV-shows through hashtags. *Multimedia Tools and Applications*, 78(3):3279–3296, 2019.

[65] Kevin M. DeLuca, Sean Lawson, and Ye Sun. Occupy Wall Street on the public screens of social media: The many framings of the birth of a protest movement. *Communication, Culture & Critique*, 5(4):483–509, 2012.

[66] Michael A. DeVito, Darren Gergle, and Jeremy Birnholtz. "algorithms ruin everything" #riptwitter, folk theories, and resistance to algorithmic change in social media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 3163–3174, 2017.

[67] D.g.f.ct.r..info. Tweet: D.g.f.ct.r..info. `https://twitter.com/dgfctr/status/806118382408003585`, 2016. [Online; Accessed 12-January-2020].

[68] Roger Dickinson, Anne Murcott, Jane Eldridge, and Simon Leader. Breakfast, time, and “breakfast time" television, food, and the household organization of consumption. *Television & New Media*, 2(3):235–256, 2001.

[69] Carl DiSalvo. Design and the construction of publics. *Design issues*, 25(1):48–63, 2009.

[70] Carl DiSalvo. *Adversarial design*. The MIT Press, 2012.

[71] Cory Doctorow. Automated calls, fraud and the banks: a mismatch made in hell. `https://www.theguardian.com/technology/2012/sep/27/automated-calls-fraud-banks`, 2012. [Online; Accessed 20-December-2019].

[72] Brian Doherty. *Radicals for capitalism: A freewheeling history of the modern American libertarian movement*. PublicAffairs, 2009.

[73] Tanja Döring, Axel Sylvester, and Albrecht Schmidt. A design space for ephemeral user interfaces. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, pages 75–82, 2013.

[74] Mark Doughty, Shaun Lawson, Conor Linehan, Duncan Rowland, and Lucy Bennett. Disinhibited abuse of othered communities by second-screening audiences. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, pages 55–62. ACM, 2014.

[75] Mark Doughty, Duncan Rowland, and Shaun Lawson. Co-viewing live TV with digital backchannel streams. In *Proceedings of the 9th European Conference on Interactive TV and Video*, pages 141–144. ACM, 2011.

[76] Paul Dourish. Implications for design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 541–550. ACM, 2006.

[77] Abigail Durrant and David Kirk. On ethical responsiveness: Being answerable to others as an HCI researcher. *Interacting with Computers*, 30(2):99–115, 2018.

[78] Leigh H. Edwards. *The triumph of reality TV: The revolution in American television*. ABC-CLIO, 2013.

[79] Robert Elliott and Ladislav Timulak. Descriptive and interpretive approaches to qualitative research. *A handbook of research methods for clinical and health psychology*, 1(7):147–159, 2005.

[80] Jon Elster. *Deliberative Democracy*. Cambridge University Press, 1998.

[81] Engineering and Physical Sciences Research Council. Research in the wild 2012. `https://epsrc.ukri.org/funding/calls/researchinthewild/`, 2012. [Online; accessed 5-November-2019].

[82] Darrell Etherington. Amazon Echo could soon start talking to you unprompted. `https://techcrunch.com/2016/09/08/amazon-echo-could-soon-start-talking-to-you-unprompted/`, 2016. [Online; Accessed 11-January-2020].

[83] Nathaniel J. Evans, Joe Phua, Jay Lim, and Hyoyeun Jun. Disclosing Instagram influencer advertising: The effects of disclosure language on advertising recognition, attitudes, and behavioral intent. *Journal of Interactive Advertising*, 17(2):138–149, 2017.

[84] Tom Feltwell, John Vines, Karen Salt, Mark Blythe, Ben Kirman, Julie Barnett, Phillip Brooker, and Shaun Lawson. Counter-discourse activism on social media: The case of challenging “poverty porn” television. *Computer Supported Cooperative Work (CSCW)*, 26(3):345–385, Jun 2017.

[85] Michelle Fitzsimmons. Alexa scares some users with unsolicited laughter, but Amazon is working on a fix. `https://www.techradar.com/news/alexa-is-laughing-at-people-unprompted-and-its-just-as-scary-as-it-sounds`, 2018. [Online; Accessed 11-January-2020].

[86] Involve NHS National Institute for Health Research. Payment for involvement. `http://www.invo.org.uk/wp-content/uploads/2012/11/INVOLVEPayment-Guiderev2012.pdf`, 2012. [Online; Accessed 01-September-2017].

[87] Michelle Forelle, Philip N. Howard, Andrés Monroy-Hernández, and Saiph Savage. Political bots and the manipulation of public opinion in Venezuela. *arXiv preprint arXiv:1507.07109*, 2015.

[88] Michel Foucault. *History of madness*. Routledge, 2013.

[89] aline shakti franzke, Anja Bechmann, Michael Zimmer, Charles Ess, and The Association of Internet Researchers. Internet research: Ethical guidelines 3.0. *Available online: https://aoir.org/reports/ethics3.pdf*, 2019.

[90] Freeview. Freeview. Video, 2018.

[91] Wai-Tat Fu, Thomas Kannampallil, Ruogu Kang, and Jibo He. Semantic imitation in social tagging. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 17(3):12, 2010.

[92] Christian Fuchs. *Social media: A critical introduction*. Sage, 2017.

[93] Roberto Gargarella. Full representation, deliberation, and impartiality. *Deliberative Democracy*, 1:260, 1998.

[94] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Reducing controversy by connecting opposing views. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 81–90. ACM, 2017.

[95] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference*, pages 913–922. International World Wide Web Conferences Steering Committee, 2018.

[96] R. Kelly Garrett. Echo chambers online?: Politically motivated selective exposure among internet news users. *Journal of Computer-Mediated Communication*, 14(2):265–285, 2009.

[97] R. Kelly Garrett. Politically motivated reinforcement seeking: Reframing the selective exposure debate. *Journal of communication*, 59(4):676–699, 2009.

[98] R. Kelly Garrett, Dustin Carnahan, and Emily K. Lynch. A turn toward avoidance? selective exposure to online political information, 2004–2008. *Political Behavior*, 35(1):113–134, 2013.

[99] Bill Gaver, Tony Dunne, and Elena Pacenti. Design: cultural probes. *interactions*, 6(1):21–29, 1999.

[100] William Gaver, Andy Boucher, Nadine Jarvis, David Cameron, Mark Hauenstein, Sarah Pennington, John Bowers, James Pike, Robin Beitra, and Liliana Ovalle. The datacatcher: batch deployment and documentation of 130 location-aware, mobile devices that put sociopolitically-relevant big data in people's hands: polyphonic interpretation at scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 1597–1607. ACM, 2016.

[101] William Gaver, Mike Michael, Tobie Kerridge, Alex Wilkie, Andy Boucher, Liliana Ovalle, and Matthew Plummer-Fernandez. Energy Babble: Mixing environmentally-oriented internet content to engage community groups. In

Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pages 1115–1124. ACM, 2015.

[102] William W Gaver, Andrew Boucher, Sarah Pennington, and Brendan Walker. Cultural probes and the value of uncertainty. *interactions*, 11(5):53–56, 2004.

[103] Matthew Gentzkow and Jesse M. Shapiro. Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126(4):1799–1839, 2011.

[104] Stella George. From sex and therapy bots to virtual assistants and tutors: how emotional should artificially intelligent agents be? In *Proceedings of the 1st International Conference on Conversational User Interfaces*, pages 1–3, 2019.

[105] Fabio Giglietto and Yenn Lee. To be or not to be charlie: Twitter hashtags as a discourse and counter-discourse in the aftermath of the 2015 Charlie Hebdo shooting in France. *AoIR Selected Papers of Internet Research*, 5, 2015.

[106] Homero Gil de Zúñiga, Victor Garcia-Perdomo, and Shannon C. McGregor. What is second screening? exploring motivations of second screen use and its effect on online political participation. *Journal of Communication*, 65(5):793–815, 2015.

[107] Katerina Gorkovenko, Nick Taylor, and Jon Rogers. Social printers: a physical social network for political debates. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2269–2281. ACM, 2017.

[108] Todd Steven Graham. *What's Wife Swap got to do with it?: Talking Politics in the net-based public sphere*. PhD thesis, Universiteit van Amsterdam [Host], 2009.

[109] Jürgen Habermas. Political communication in media society: Does democracy still enjoy an epistemic dimension? the impact of normative theory on empirical research. *Communication theory*, 16(4):411–426, 2006.

[110] Jurgen Habermas and Jürgen Habermas. *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. MIT press, 1991.

[111] Asmelash Teka Hadgu, Kiran Garimella, and Ingmar Weber. Political hashtag hijacking in the us. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 55–56, 2013.

[112] Stuart Hall. Who needs identity. *Questions of Cultural Identity*, 16(2):1–17, 1996.

[113] Rahaf Harfoush. *Yes We Did! An inside look at how social media built the Obama brand*. New Riders, 2009.

[114] Eszter Hargittai, Jason Gallo, and Matthew Kane. Cross-ideological discussions among conservative and liberal bloggers. *Public Choice*, 134(1-2):67–86, 2008.

[115] Richard Harper, Tom Rodden, Yvonne Rogers, and Abigail Sellen. *Being Human: Human-Computer Interaction In The Year 2020*. Microsoft Research Ltd, 2008.

[116] Gerard A. Hauser and Erin Diana McClellan. Vernacular rhetoric and social movements: Performances of resistance in the rhetoric of the everyday. *Active voices: Composing a rhetoric of social movements*, pages 23–46, 2009.

[117] Roberta Hawkins and Jennifer J. Silver. From selfie to #sealfie: Nature 2.0 and the digital cultural politics of an internationally contested resource. *Geoforum*, 79:114–123, 2017.

[118] HBO. Game of Thrones interactive experience. Website: `http://www.hbogo.com/geo.html`, or `http://viewers-guide.hbo.com/`, 2013.

[119] Samuel Hooton. Social media echo chambers gifted donald trump the presidency. `https://www.independent.co.uk/voices/donald-trump-president-social-media-echo-chamber-hypernormalisation-adam-curtis-protests-blame-a7409481.html`, 2016. [Online; Accessed 25-February-2020].

[120] Philip N. Howard, Bence Kollanyi, and Samuel Woolley. Bots and automation over Twitter during the US election. *Computational Propaganda Project: Working Paper Series*, 2016.

[121] Jacob H. Huebert. *Libertarianism today*. ABC-CLIO, 2010.

[122] Zaheer Hussain and Mark D. Griffiths. The attitudes, feelings, and experiences of online gamers: a qualitative analysis. *CyberPsychology & Behavior*, 12(6):747–753, 2009.

[123] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 17–24, 2003.

[124] Jelani Ince, Fabio Rojas, and Clayton A. Davis. The social media response to Black Lives Matter: how Twitter users interact with Black Lives Matter through hashtag use. *Ethnic and Racial studies*, 40(11):1814–1830, 2017.

[125] Shanto Iyengar. *Is anyone responsible?: How television frames political issues*. University of Chicago Press, 1994.

[126] Henry Jenkins, Mizuko Ito, and danah boyd. *Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics*. John Wiley & Sons, 2015.

[127] Tracey Jensen. Welfare commonsense, poverty porn and doxosophy. *Sociological Research Online*, 19(3):1–7, 2014.

[128] Qihao Ji and Arthur A. Raney. Morally judging entertainment: A case study of live tweeting during Downton Abbey. *Media Psychology*, 18(2):221–242, 2015.

[129] Owen Jones. Owen jones: My latest battle to stop the demonising of people on benefits. `https://www.independent.co.uk/voices/comment/owen-jones-my-latest-battle-to-stop-the-demonising-of-people-on-benefits-9136123.html`, 2014. [Online; Accessed 10-February-2020].

[130] Yvonne Kammerer, Rowan Nairn, Peter Pirolli, and Ed H. Chi. Signpost from the masses: learning effects in an exploratory social tag search browser. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 625–634. ACM, 2009.

[131] Habibul Haque Khondker. Role of the new media in the Arab Spring. *Globalizations*, 8(5):675–679, 2011.

[132] Ben Kirman, Conor Linehan, and Shaun Lawson. Get lost: facilitating serendipitous exploration in location-sharing services. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 2303–2308. ACM, 2012.

[133] Ben Kirman, Conor Linehan, Shaun Lawson, Derek Foster, and Mark Doughty. There's a monster in my kitchen: using aversive feedback to motivate behaviour change. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pages 2685–2694. ACM, 2010.

[134] Lorenz Cuno Klopfenstein, Saverio Delpriori, Silvia Malatini, and Alessandro Bogliolo. The rise of bots: A survey of conversational interfaces, patterns, and paradigms. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, pages 555–565. ACM, 2017.

[135] Matthias Korn and Amy Voida. Creating friction: infrastructuring civic engagement in everyday life. In *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives*, pages 145–156. Aarhus University Press, 2015.

[136] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790, 2014.

[137] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. Supporting reflective public thought with ConsiderIt. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, pages 265–274. ACM, 2012.

[138] Laura Kurgan. Million-dollar blocks. 2013.

[139] Ben Lamb. Cathy come off benefits: A comparative ideological analysis of Cathy Come Home and Benefits Street. *Journalism and Discourse Studies*, 2016.

[140] David R. Large, Leigh Clark, Gary Burnett, Kyle Harrington, Jacob Luton, Peter Thomas, and Pete Bennett. “it’s small talk, Jim, but not as we know it." engendering trust through human-agent conversation in an autonomous, self-driving car. In *Proceedings of the 1st International Conference on Conversational User Interfaces*, pages 1–7, 2019.

[141] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):102, 2018.

[142] Effie Le Moignan, Shaun Lawson, Duncan A. Rowland, Jamie Mahoney, and Pam Briggs. Has Instagram fundamentally altered the ‘family snapshot’? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 4935–4947. ACM, 2017.

[143] Jyh-An Lee and Ching-U Liu. Forbidden City enclosed by the Great Firewall: The law and power of Internet filtering in China. *Minn. JL Sci. & Tech.*, 13:125, 2012.

[144] Conor Linehan, Sabine Harrer, Ben Kirman, Shaun Lawson, and Marcus Carter. Games against health: A player-centered design philosophy. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’15, pages 589–600, New York, NY, USA, 2015. ACM.

[145] Eden Litt and Eszter Hargittai. The imagined audience on social network sites. *Social Media+ Society*, 2(1):2056305116633482, 2016.

[146] Bjorn Lomborg. *Cool it: the skeptical environmentalist’s guide to global warming*. Vintage, 2007.

[147] Gilad Lotan, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce, and danah boyd. The Arab Spring | the revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *International journal of communication*, 5:31, 2011.

[148] James Lull. The social uses of television. *Human communication research*, 6(3):197–209, 1980.

[149] Karen Lumsden and Heather Morgan. Media framing of trolling and online abuse: silencing strategies, symbolic violence, and victim blaming. *Feminist Media Studies*, 17(6):926–940, 2017.

[150] Robert MacDonald, Tracy Shildrick, and Andy Furlong. 'Benefits Street' and the myth of workless communities. *Sociological Research Online*, 19(3):1–6, 2014.

[151] Jamie Mahoney, Tom Feltwell, Obinna Ajuruchi, and Shaun Lawson. Constructing the visual online political self: an analysis of Instagram use by the Scottish electorate. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 3339–3351. ACM, 2016.

[152] Jane Mansbridge. Everyday talk in the deliberative system. 1999.

[153] Nathalie Maréchal. When bots tweet: Toward a normative framework for bots on social networking sites. *International Journal of Communication*, 10:10, 2016.

[154] Annette Markham and Elizabeth Buchanan. Ethical decision-making and internet research: Version 2.0. recommendations from the AoIR ethics working committee. *Available online: aoir. org/reports/ethics2. pdf*, 2012.

[155] Maxwell E. McCombs and Donald L. Shaw. The agenda-setting function of mass media. *Public opinion quarterly*, 36(2):176–187, 1972.

[156] Miller McPherson, Lynn Smith-Lovin, and James M. Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444, 2001.

[157] Joshua McVeigh-Schultz, Anya Kolesnichenko, and Katherine Isbister. Shaping pro-social interaction in VR: An emerging design framework. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

[158] Nigel Meek. Personal and economic ideology: British party politics and the political compass. *Political Notes*, 155:1–51, 1999.

[159] Yelena Mejova, Sofiane Abbar, and Hamed Haddadi. Fetishizing food in digital age: #foodporn around the world. In *Tenth International AAAI Conference on Web and Social Media*, 2016.

[160] Donald Melanson. iPhone push notification service for devs announced. `https://www.engadget.com/2008/06/09/iphone-push-notification-service-for-devs-announced/`, 2008. [Online; Accessed 11-January-2020].

[161] Jonathan Mellon and Christopher Prosser. Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics*, 4(3):2053168017720008, 2017.

[162] Abraham Mhaidli, Manikandan Kandadai Venkatesh, Yixin Zou, and Florian Schaub. Listen only when spoken to: Interpersonal communication cues as smart speaker privacy controls. *Proceedings on Privacy Enhancing Technologies*, 1:20, 2019.

[163] I-Spy (Michelin). I-spy (michelin). `https://en.wikipedia.org/wiki/I-Spy_(Michelin)`, 2019. [Online; Accessed 21-March-2019].

[164] Carl Miller. *The death of the gods: The new global power grab*. Random House, 2018.

[165] Bill Moggridge and Bill Atkinson. *Designing interactions*, volume 17. MIT press Cambridge, MA, 2007.

[166] Shaheed N. Mohammed. Conspiracy theories and flat-earth videos on youtube. *The Journal of Social Media in Society*, 8(2):84–102, 2019.

[167] Evgeny Morozov. *To save everything, click here: The folly of technological solutionism*. Public Affairs, 2013.

[168] Chantal Mouffe. *The democratic paradox*. Verso Books, 2000.

[169] Chantal Mouffe. *Agonistics: Thinking the world politically*. Verso Books, 2013.

[170] Allison Mountz. The Other. In *Key Concepts in Political Geography*, pages 328–339. SAGE, 2009.

[171] Sean A. Munson, Stephanie Y. Lee, and Paul Resnick. Encouraging reading of diverse political viewpoints with a browser widget. In *Seventh International AAAI Conference on Weblogs and Social Media*, 2013.

[172] Sean A. Munson and Paul Resnick. Presenting diverse political opinions: how and how much. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1457–1466. ACM, 2010.

[173] Sean A. Munson and Paul Resnick. The prevalence of political discourse in non-political blogs. In *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[174] Janet Murray, Sergio Goldenberg, Kartik Agarwal, Tarun Chakravorty, Jonathan Cutrell, Abraham Doris-Down, and Harish Kothandaraman. Storymap: iPad companion for long form TV narratives. In *Proceedings of the 10th European Conference on Interactive TV and Video*, pages 223–226. ACM, 2012.

[175] Diana C. Mutz. *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press, 2006.

[176] Abhishek Nandakumar and Janet Murray. Companion apps for long arc TV series: supporting new viewers in complex storyworlds with tightly synchronized context-sensitive annotations. In *Proceedings of the ACM International*

Conference on Interactive Experiences for TV and Online Video, pages 3–10. ACM, 2014.

[177] Kate Nash. *Contemporary political sociology: Globalization, politics and power*. John Wiley & Sons, 2009.

[178] Nir Noon Nave, Limor Shifman, and Keren Tenenboim-Weinblatt. Talking it personally: Features of successful political posts on Facebook. *Social Media+ Society*, 4(3):2056305118784771, 2018.

[179] Les Nelson, Christoph Held, Peter Pirolli, Lichan Hong, Diane Schiano, and Ed H. Chi. With a little help from my friends: examining the impact of social annotations in sensemaking tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1795–1798. ACM, 2009.

[180] Azadeh Nematzadeh, Giovanni Luca Ciampaglia, Yong-Yeol Ahn, and Alessandro Flammini. Information overload in group communication: From conversation to cacophony in the Twitch chat. *Royal Society Open Science*, 6(10):191412, 2016.

[181] Mai-Hanh Nguyen. From fortune 500s to small businesses, real businesses are already using chatbots to improve their service. `https://www.businessinsider.com/examples-brands-companies-chatbots-business-2017-10`, 2017. [Online; Accessed 20-December-2019].

[182] Safiya Umoja Noble. *Algorithms of oppression: How search engines reinforce racism*. nyu Press, 2018.

[183] David Nolan. Classifying and analysing politico-economic systems. *The Individualist*, 1:5–11, 1971.

[184] Brendan Nyhan and Jason Reifler. When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2):303–330, 2010.

[185] House of Commons. Topical Questions - House of Commons Hansard. https://hansard.parliament.uk/Commons/2014-01-13/debates/14011313000027/TopicalQuestions, 2014. [Online; Accessed 10-February-2020].

[186] Chong Oh, Sheila Sasser, and Soliman Almahmoud. Social media analytics framework: the case of twitter and super bowl ads. *Journal of Information Technology Management*, 26(1):1–18, 2015.

[187] Ray Oldenburg. *The great good place: Café, coffee shops, community centers, beauty parlors, general stores, bars, hangouts, and how they get you through the day*. Paragon House Publishers, 1989.

[188] Cathy O'Neil. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books, 2016.

[189] Walter J. Ong. *Orality and literacy*. Routledge, 2013.

[190] Anthony M Orum. *Introduction to political sociology: The social anatomy of the body politic*. Prentice-Hall Englewood Cliffs, NJ, 1978.

[191] Eli Pariser. *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011.

[192] Sine Plambech. *Terror in New Zealand: a lesson in Othering the terrorists, not the migrants!* DIIS, 2019.

[193] John Plunkett. Ralph Lee: 'we don't have a malign intention, there is no agenda'. https://www.theguardian.com/media/2014/oct/26/channel-4-ralph-lee-benefits-street-immigration, 2012. [Online; Accessed 01-September-2017].

[194] Bailey Poland. *Haters: Harassment, abuse, and violence online*. University of Nebraska Press, 2016.

[195] Martin Porcheron, Joel E. Fischer, Stuart Reeves, and Sarah Sharples. Voice interfaces in everyday life. In *proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 640. ACM, 2018.

[196] Martin Porcheron, Joel E. Fischer, and Sarah Sharples. Using mobile phones in pub talk. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 1649–1661, 2016.

[197] john a. powell. Us vs them: the sinister techniques of 'othering' - and how to avoid them. https://www.theguardian.com/inequality/2017/nov/08/us-vs-them-the-sinister-techniques-of-othering-and-how-to-avoid-them, 2017. [Online; Accessed 31-January-2020].

[198] Jennifer Preece, Yvonne Rogers, and Helen Sharp. *Interaction design*. Wiley, 2019.

[199] Marguerite Reardon. Your alexa and fitbit can testify against you in court. https://www.cnet.com/news/alexa-fitbit-apple-watch-pacemaker-can-testify-against-you-in-court/, 2018. [Online; Accessed 27-February-2020].

[200] Stuart Reeves. Conversation considered harmful? In *Proceedings of the 1st International Conference on Conversational User Interfaces*, pages 1–3, 2019.

[201] Stuart Reeves, Martin Porcheron, and Joel E. Fischer. 'this is not what we wanted': designing for conversation with voice interfaces. *interactions*, 26(1):46–51, 2018.

[202] EU Regulation. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, pages L119/1, p. 1–88, 2016.

[203] Alastair Reid and Carlotta Dott. Thousands of misleading conservative ads side-step scrutiny thanks to facebook policy. https://firstdraftnews.org/latest/thousands-of-misleading-conservative-ads-side-step-scrutiny-thanks-to-facebook-policy/, 2019. [Online; Accessed 21-January-2020].

[204] Matt Reynolds. Microsoft is betting that bots 'are the new apps'. https://www.wired.co.uk/article/microsoft-build-bots-ai-cortana-keynote-conference, 2016. [Online; Accessed 20-December-2019].

[205] Juliet Richardson. Are focus groups a useful research technique in UX? https://www.nomensa.com/blog/2016/are-focus-groups-useful-research-technique-ux, 2016. [Online; accessed 01-November-2019].

[206] Stephen Harold Riggins. The rhetoric of Othering. *The language and politics of exclusion: Others in discourse*, 8:1–30, 1997.

[207] Colin Robson. *Real world research*, volume 3. Wiley Chichester, 2011.

[208] Kathleen Rodgers and Willow Scobie. Sealfies, seals and celebs: expressions of Inuit resilience in the Twitter era. *Interface: A Journal on Social Movements*, 7(1):70–97, 2015.

[209] Yvonne Rogers. Interaction design gone wild: striving for wild theory. *interactions*, 18(4):58–62, 2011.

[210] Jon Ronson. *So you've been publicly shamed*. Riverhead Books, 2016.

[211] John Rooksby. Wild in the laboratory: a discussion of plans and situated actions. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(3):19, 2013.

[212] Carl Sagan. *The demon-haunted world: Science as a candle in the dark*. Ballantine Books, 2011.

[213] Edward W. Said. *Orientalism*. Pantheon Books, 1978.

[214] Johnny Saldana. *Thinking qualitatively: Methods of mind*. Sage Publications, 2014.

[215] B. N. Sanders, Elizabeth and P Stappers. *Convivial toolbox: generative research for the front end of design*. BIS Publishers, 2012.

[216] Cornel Sandvoss, Kelly Youngs, and Joanne Hobbs. Television fandom in the age of narrowcasting: The politics of proximity in regional scripted reality dramas The Only Way Is Essex and Made in Chelsea. In *Popular Media Cultures*, pages 39–72. Springer, 2015.

[217] Roscoe C. Scarborough and Charles Allan McCoy. Moral reactions to reality TV: Television viewers' endogenous and exogenous loci of morality. *Journal of Consumer Culture*, 16(1):164–191, 2016.

[218] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I. Hong. Hey Alexa, what's up?: A mixed-methods studies of in-home conversational agent usage. In *Proceedings of the 2018 Designing Interactive Systems Conference*, pages 857–868. ACM, 2018.

[219] Bryan C. Semaan, Scott P. Robertson, Sara Douglas, and Misa Maruyama. Social media supporting political deliberation across multiple public spheres: towards depolarization. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 1409–1421. ACM, 2014.

[220] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. Reflective design. In *Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility*, CC '05, pages 49–58, New York, NY, USA, 2005. ACM.

[221] Tracy Shildrick, Robert MacDonald, and Andy Furlong. 'Benefit street' & the myth of workless communities. *Sociological Research Online*, 19(3):1–6, 2014.

[222] British Psychological Society. Ethics guidelines for internet-mediated research. `https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20-%20Files/Ethics%20Guidelines%20for%20Internet-mediated%20Research%20%282017%29.pdf`, 2017. [Online; Accessed 21-March-2019].

[223] Dominic Spohr. Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*, 34(3):150–160, 2017.

[224] Fiona Ssozi-Mugarura, Thomas Reitmaier, Anja Venter, and Edwin Blake. Enough with “In-The-Wild". In *Proceedings of the First African Conference on Human Computer Interaction*, AfriCHI'16, pages 182–186, New York, NY, USA, 2016. Association for Computing Machinery.

[225] Susan Stokes. Pathologies of deliberation. *Deliberative democracy*, 1:123, 1998.

[226] John Suler. The online disinhibition effect. *Cyberpsychology & behavior*, 7(3):321–326, 2004.

[227] Cass R. Sunstein. The ethics of nudging. *Yale Journal on Regulation*, 32:413, 2015.

[228] Selina Jeanne Sutton, Paul Foulkes, David Kirk, and Shaun Lawson. Voice as a design material: Sociophonetic inspired design strategies in human-computer interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 603. ACM, 2019.

[229] Jonathan Taplin. *Move fast and break things: How Facebook, Google, and Amazon have cornered culture and what it means for all of us*. Pan Macmillan, 2017.

[230] Gareth Terry, Nikki Hayfield, Victoria Clarke, and Virginia Braun. Thematic analysis. *The Sage handbook of qualitative research in psychology*, pages 17–37, 2017.

[231] Yannis Theocharis, Will Lowe, Jan W. van Deth, and Gema García-Albacete. Using Twitter to mobilize protest action: online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. *Information, Communication & Society*, 18(2):202–220, 2015.

[232] Jennifer Thom-Santelli, Michael J. Muller, and David R. Millen. Social tagging roles: publishers, evangelists, leaders. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1041–1044. ACM, 2008.

[233] Vanessa Thomas. Tweet thread: Vanessa Thomas Ph.D on Twitter. https://twitter.com/drt3ch/status/805375626404331522, 2016. [Online; Accessed 12-January-2020].

[234] Clive Thompson. *Coders: The Making of a New Tribe and the Remaking of the World*. Penguin, 2019.

[235] Peter Tolmie. Orienting to the wild. In *Into the Wild: Beyond the Design Research Lab*, pages 195–235. Springer, 2020.

[236] Annabel Tremlett. Demotic or demonic? race, class and gender in 'Gypsy' reality TV. *The Sociological Review*, 62(2):316–334, 2014.

[237] Cristian Vaccari, Andrew Chadwick, and Ben O'Loughlin. Dual screening the political: Media events, social media, and citizen engagement. *Journal of Communication*, 65(6):1041–1061, 2015.

[238] Sebastián Valenzuela, Martina Piña, and Josefina Ramírez. Behavioral effects of framing on social media users: How conflict, economic, human interest, and morality frames drive news sharing. *Journal of Communication*, 67(5):803–826, 2017.

[239] José Van Dijck, Thomas Poell, and Martijn De Waal. *The platform society: Public values in a connective world*. Oxford University Press, 2018.

[240] Karin van Es, Daniela Van Geenen, and Thomas Boeschoten. Mediating the black pete discussion on Facebook: Slacktivism, flaming wars, and deliberation. *First Monday*, 19(12), 2014.

[241] Bert Vandenberghe and Karin Slegers. Designing for others, and the trap of HCI methods & practices. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '16, pages 512–524, New York, NY, USA, 2016. ACM.

[242] James Vincent. Google's AI sounds like a human on the phone - should we be worried? https://www.theverge.com/2018/5/9/17334658/google-ai-phone-call-assistant-duplex-ethical-social-implications, 2018. [Online; Accessed 27-January-2020].

[243] Vasilis Vlachokyriakos, Rob Comber, Karim Ladha, Nick Taylor, Paul Dunphy, Patrick McCorry, and Patrick Olivier. PosterVote: expanding the action repertoire for local political activism. In *Proceedings of the 2014 Conference on Designing Interactive Systems*, pages 795–804. ACM, 2014.

[244] Vasilis Vlachokyriakos, Paul Dunphy, Nick Taylor, Rob Comber, and Patrick Olivier. BallotShare: An exploration of the design space for digital voting in the workplace. *Computers in Human Behavior*, 41:433–443, 2014.

[245] Alex Hai Wang. Detecting spam bots in online social networking sites: a machine learning approach. In *IFIP Annual Conference on Data and Applications Security and Privacy*, pages 335–342. Springer, 2010.

[246] Christina Warren. Facebook finally gets #hashtags. `https://edition.cnn.com/2013/06/12/tech/social-media/facebook-hashtags/index.html`, 2013. [Online; Accessed 13-January-2020].

[247] Max Weber. *Economy and society: An outline of interpretive sociology*, volume 1. University of California Press, 1978.

[248] Joseph Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45, 1966.

[249] Rolf Wiesemes. Academics teaching controversial and sensitive issues in schools: Transferring findings from genocide education research into outreach practice. *Arts and Humanities Academics in Schools: Mapping the Pedagogical Interface*, pages 142–150, 2011.

[250] Carla Willig. *Introducing Qualitative Research in Psychology*. Sage, 2008.

[251] Gavin Wood, Kiel Long, Tom Feltwell, Scarlett Rowland, Phillip Brooker, Jamie Mahoney, John Vines, Julie Barnett, and Shaun Lawson. Rethinking engagement with online news through social and visual co-annotation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 576. ACM, 2018.

[252] Thomas Wood and Ethan Porter. The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior*, 41(1):135–163, 2019.

[253] Scott Wright, Todd Graham, and Dan Jackson. Third space, social media, and everyday political talk. In *The Routledge companion to social media and politics*, pages 74–88. Routledge, 2015.

[254] Scott Wright and John Street. Democracy, deliberation and design: the case of online discussion forums. *New Media & Society*, 9(5):849–869, 2007.

[255] Bin Xu, Pamara Chang, Christopher L. Welker, Natalya N. Bazarova, and Dan Cosley. Automatic archiving versus default deletion: what Snapchat tells us about ephemerality in design. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 1662–1675. ACM, 2016.

[256] Slavoj Žižek. Denial: The liberal utopia. *Lacan. com*, 2009.